LECTURE NOTES ON QUANTUM MECHANICS

PHY 851/852 – 2020/2021/2022

Scott Pratt

Department of Physics and Astronomy
Michigan State University

FORWARD

These notes are for the two-semester graduate level quantum mechanics class taught at Michigan State University. Although they are more terse than a typical text book, they do cover all the material used in PHY 851/852. The notes presume a familiarity with basic undergraduate concepts in quantum mechanics.

Anybody is welcome to use the notes to their heart's content, though the text should be treated with the usual academic respect when it comes to copying material. If anyone is interested in the LaTeX source files, they should contact me (prattsc@msu.edu). Solutions to the end-of-chapter problems will also be provided on the course web site, https://people.nscl.msu.edu/~pratt/phy851. Please beware that this is a web manuscript, and is thus a living document and subject to change at any time.

Contents

1 States and Operators

1.1 States

Quantum mechanics consists of states and operators. For any finite system, states are discrete and can be assigned labels. The discrete nature of states is what is what is meant by the word "quantized". For some systems the number of such discrete states is finite, e.g. a particle in either spin-up or spin-down state, whereas in other systems the number of states are infinite, e.g. levels of a harmonic oscillator. In this chapter we concentrate on those systems where the number of states are finite, and proceed to the latter case in the next chapter.

Any physical state can be assigned a label, in this case ψ. This state can always be expressed as a linear combination of basis states, $|i\rangle$, in an orthonormal basis. These states can be expressed as vectors in a vector space of size n,

$$|1\rangle = \hat{1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad |2\rangle = \hat{2} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \cdots |n\rangle = \hat{n} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \tag{1.1}$$

By inspection, the basis is orthonormal,

$$\langle i|j\rangle \equiv \hat{i}^* \cdot \hat{j} = \delta_{ij}. \tag{1.2}$$

In quantum mechanics the vector algebra is complex, and the adjoint vector $\langle\psi|$ is represented by the complex transpose. Any orthonormal set of vectors can be expressed in this basis. The notational choice of using $\langle\psi|$ to refer to the adjoint vector and $|\psi\rangle$ to denote the vector is known as *bra*s and *ket*s respectively (a take on the word "bracket"), and is known as Dirac notation, though it derives from Hermann Grassmann's work, https://en.wikipedia.org/wiki/Hermann_Grassmann, one hundred years earlier.

Whenever the notation has a bra followed by a ket, it implies the dot product between the two vectors, i.e.

$$\langle v|u\rangle = \sum_i v_i^* u_i, \tag{1.3}$$

and is sometimes referred to as an *inner* product. A sum over vector indices is not implied if the ket precedes the bra, and

$$|u\rangle\langle v| = u_i v_j^* \tag{1.4}$$

is known as an outer product. However, if one sees the notation,

$$|\alpha\rangle\langle\alpha|, \tag{1.5}$$

a sum over the vectors α is implied (unless otherwise stated), as the summation $\sum_\alpha$ is implicit. Ultimately, for any physical observable, all bras and kets will be closed with a bra being the first

to the left and a ket being the last to the right, with all vector indices being summed over within each bra-ket factor.

One can form new states by taking linear combinations of basis states $|i\rangle$. For example,

$$|\psi\rangle = \sum_i a_{\psi,i}|i\rangle$$
$$\langle\psi| = \sum_i a^*_{\psi,i}\langle i|,$$

where $a_{i,\psi}$ is a complex number. By taking the overlap of $|\psi\rangle$ with $\langle i|$, one sees that

$$a_{i,\psi} = \langle i|\Psi\rangle. \tag{1.6}$$

If the basis states $|i\rangle$ are orthonormal and complete, one can state that

$$\sum_i |i\rangle\langle i| = \mathbb{I}. \tag{1.7}$$

This is known as the completeness relation. It is easy to see that completeness relation holds for the simple basis in Eq. (1.1). Any other orthonormal basis can be related to this simple basis by a basis transformation, which is an example of a unitary transformation. A basis transformation relates a set of basis states $|i\rangle$ to a set of basis states $|j\rangle$. One can express the state $|i\rangle$ in terms of the basis j,

$$|i\rangle = \sum_j |j\rangle\langle j|i\rangle. \tag{1.8}$$

If this is to work for any state $i\rangle$, the completeness relation in Eq. (1.7) must hold.

Furthermore, by taking the overlap of ψ with itself (the norm),

$$\begin{aligned}\langle\psi|\psi\rangle &= \sum_i \langle\psi|i\rangle\langle i|\psi\rangle \\ &= \sum_i |\langle\psi|i\rangle|^2 = \sum_i |a_{\psi,i}|^2.\end{aligned} \tag{1.9}$$

If the state $|\psi\rangle$ is normalized, i.e. $\langle\psi|\psi\rangle = 1 = \sum_i |\langle i|\psi\rangle|^2 = \sum_i |a_{\psi,i}|^2$, the squared element, $|\langle i|\psi\rangle|^2$, can be interpreted as a probability because each term is positive and because the sum yields unity. Thus, if one prepares a state $|\psi\rangle$, the chance one finds that one is in a state $|i\rangle$ is $|\langle i|\psi\rangle|^2$. Up to this last sentence, where the connection to probability is stated, the discussion has been one of standard linear algebra and notation. Of course, once one is in the state i, you know that the system is in the state i, and any knowledge of being in the original state ψ is lost.

Associating the squared overlap as a probability is the profound intellectual jump that makes quantum mechanics physics. Most of the expressions applied in this course are derived, almost inexorably, from this conceptual leap combined with arguments about symmetry and the need to reproduce classical mechanics in some limit.

Equation (1.9) also suggests the completeness relation,

$$\sum_i |i\rangle\langle i| = \mathbb{I}. \tag{1.10}$$

This is easy to see with the simple basis states in Eq. (1.1). In that case

$$|1\rangle\langle 1| = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ & & \vdots & \\ 0 & & \cdots, & 0 \end{pmatrix}, \quad |2\rangle\langle 2| = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & \vdots & \\ 0 & & \cdots, & 0 \end{pmatrix}, \quad \cdots \tag{1.11}$$

The completeness relation will also work for any set of basis states, because the basis can always be transformed to the simple basis, and because $\sum_i |i\rangle\langle i| = \mathbb{I}$ and the unit matrix is unchanged by a transformation of basis, a.k.a. a unitary transformation. This will be demonstrated ahead.

As mentioned previously, all states can be expressed as vectors. For instance, if a basis has two states, the two states might be defined in terms of vectors as,

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{1.12}$$

The choice of these two vectors is arbitrary. As long as the two vectors are orthonormal, any state can be represented as a linear sum of the two states. Any set of orthonormal states that spans the space is known as a "basis", with each orthonormal state being a basis state.

Example 1.1: Probabilities and Overlaps
Consider the state

$$|\psi\rangle = \frac{1}{Z^{1/2}} \begin{pmatrix} 1 \\ i \end{pmatrix}.$$

a) Find Z so that the state is normalized.

b) What is the probability that $|\psi\rangle$ would be measured in the $|\downarrow\rangle$ state?

Solution:
a) Squaring $|\psi\rangle$,

$$\begin{aligned} \langle\psi|\psi\rangle &= \frac{1}{Z} \begin{pmatrix} 1 & -i \end{pmatrix} \begin{pmatrix} 1 \\ i \end{pmatrix} \\ &= \frac{2}{Z}, \\ Z &= 2. \end{aligned}$$

b) The probability is

$$\begin{aligned} P(\downarrow) &= |\langle\downarrow|\psi\rangle|^2 \\ &= \frac{1}{2} \left| \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ i \end{pmatrix} \right|^2 \\ &= \frac{1}{2}(-i)(i) = \frac{1}{2}. \end{aligned}$$

1.2 Operators

Operators operate on vectors and return another vector. All operators may be described in terms of bras and kets as a linear combination of outer products,

$$\mathcal{A} = \sum_{ij} a_{ij} |i\rangle\langle j|. \tag{1.13}$$

Just as any state can be equivalently expressed as a vector, any operator can be represented by a matrix, in this case a_{ij}. Just as a state's expression in terms of a vector depends on the basis, so does the expression of an operator in terms of a matrix. Knowing the coeficients a_{ij} is sufficient to define the matrix. If the basis is defined by

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \end{pmatrix}, \quad |2\rangle = \begin{pmatrix} 0 \\ 1 \\ \vdots \end{pmatrix}, \quad \cdots, \tag{1.14}$$

the matix and the coefficients are synonymous $A_{ij} = a_{ij}$. Otherwise, one can express the states $|i\rangle$ and $|j\rangle$ as vectors, $v(i)_m$ and $v(j)_n$. In that case

$$A_{mn} = \sum_{ij} a_{ij} v(i)_m v^*(j)_n, \tag{1.15}$$

where $v(i)_m$ is the m^{th} component of the basis vector $v(i)$.

1.3 Hermitian Conjugate:

Consider a matrix element, $\langle\phi|\mathcal{A}|\psi\rangle$. Expressing the operators as matrices and the states as vectors, then taking the complex conjugate one sees that,

$$\begin{aligned} \langle\phi|\mathcal{A}|\psi\rangle &= \phi_i^* a_{ij} \psi_j \\ \langle\phi|\mathcal{A}|\psi\rangle^* &= \left(\phi_i^* a_{ij} \psi_j\right)^* \\ &= \psi_j^* a_{ij}^* \phi_j. \end{aligned} \tag{1.16}$$

Thus, if one wishes to define an operator $\mathcal{A}^\dagger$ such that

$$\langle\psi|\mathcal{A}^\dagger|\phi\rangle = \langle\phi|\mathcal{A}|\psi\rangle^*, \tag{1.17}$$

for any states $|\phi\rangle$ and $|\psi\rangle$, the operator $\mathcal{A}^\dagger$ must be represented by the complex-conjugate of the matrix that represents $\mathcal{A}$, then transposed.

$$\begin{aligned} A^\dagger &= \sum_{ij} (a^\dagger)_{ij} |i\rangle\langle j|, \\ (a^\dagger)_{ij} &= a_{ji}^* \end{aligned} \tag{1.18}$$

The operator $\mathcal{A}^\dagger$ is known as the Hermitian conjugate of $\mathcal{A}$. A Hermitian operator is one that obeys the relation,

$$\mathcal{A} = \mathcal{A}^\dagger. \tag{1.19}$$

The eigenvalues of a Hermitian operator are real, i.e. one can always find a change of basis that diagonalizes $\mathcal{A}$, with the elements all being real. To see this, assume the operator $\boldsymbol{A}$ has eigenvalues a_i and eigenvectors $|i\rangle$.

$$\begin{aligned} \mathcal{A}|i\rangle &= a_i|i\rangle, \\ \langle i|\mathcal{A}^\dagger &= a_i^*\langle i|, \\ \langle i|(\mathcal{A}-\mathcal{A}^\dagger)|i\rangle &= a_i - a_i^*. \end{aligned} \tag{1.20}$$

If $\mathcal{A}$ is Hermitian, $\mathcal{A} = \mathcal{A}^\dagger$, and a_i must equal a_i^*, so the eigenvalues are real.

Hermitian operators are important for two reasons,

1. They yield real numbers as expectations, $\langle\psi|\mathcal{K}|\psi\rangle^* = \langle\psi|\mathcal{K}|\psi\rangle$. Because all physical observables are real, they need to be represented by Hermitian operators. An example is the Hamiltonian $\boldsymbol{H}$ which represents the energy of a system.
2. They can be used to generate unitary transformations, e.g. $e^{i\boldsymbol{K\theta}}$. An example is the time evolution operator $e^{-i\boldsymbol{Ht}}$. This will be discussed briefly in the next sub-section, but plays a much larger role when we have a general discussion of symmetries.

1.4 Unitary Operators and Transformations

A class of operators that often play an important physical role is defined by those that correspond to a change of basis, while keeping the new basis orthonormal. Examples are rotations, translations and reflections. A common unitary operator is the evolution operator, $\boldsymbol{U} = e^{-i\boldsymbol{H}t/\hbar}$, which enacts translations in time. Consider two sets of basis states $|i\rangle$ and $|i'\rangle$, related by a transformation operator $\mathcal{R}$,

$$\begin{aligned} |i'\rangle &= \sum_i \boldsymbol{R}_{i'i}|i\rangle, \\ \langle i'| &= \sum_i \boldsymbol{R}^*_{i'i}\langle i| \\ &= \sum_i \langle i|\boldsymbol{R}^\dagger_{ii'}. \end{aligned} \tag{1.21}$$

Here, the set of states $|i'\rangle$ will refer to a new basis. The coefficients $\mathcal{R}_{i'i}$ must be constructed to retain the orthormality properties,

$$\begin{aligned} \langle i'|j'\rangle &= \delta_{i'j'} \\ &= \sum_{ij} \boldsymbol{R}^\dagger_{ii'}\boldsymbol{R}_{j'j}\langle i|j\rangle \\ &= \sum_i \boldsymbol{R}_{j'i}\boldsymbol{R}^\dagger_{ii'} \\ &= (\boldsymbol{R}\boldsymbol{R}^\dagger)_{j'i'}. \end{aligned} \tag{1.22}$$

Here, the Hermitian conjugate $\mathcal{R}^\dagger$ was defined in Eq. (1.18). In simple words, the Hermitian conjugate of a unitary matrix is its inverse, i.e., $\mathcal{R}\mathcal{R}^\dagger = \mathbb{I}$, or equivalently, $\boldsymbol{R}^{-1} = \boldsymbol{R}^\dagger$. An

operator that satisfies this condition is called *unitary*. The word unitary follows from the fact that a state $\boldsymbol{U}|\psi\rangle$ has the same norm as $|\psi\rangle$, implying that the net probability is unchanged by the unitary transformation $\boldsymbol{U}$.

Hermitian operators are often used to generate unitary transformations,

$$\boldsymbol{U} = \boldsymbol{e}^{-i\mathcal{K}\boldsymbol{\theta}}. \tag{1.23}$$

It is easy to see that such an operator is unitary if $\mathcal{K} = \mathcal{K}^\dagger$,

$$\begin{aligned} \boldsymbol{U}\boldsymbol{U}^\dagger &= \boldsymbol{e}^{-i\mathcal{K}\boldsymbol{\theta}}\boldsymbol{e}^{i\mathcal{K}^\dagger\boldsymbol{\theta}} \\ &= \boldsymbol{e}^{-i\mathcal{K}\boldsymbol{\theta}}\boldsymbol{e}^{i\mathcal{K}\boldsymbol{\theta}} = \boldsymbol{e}^{-i(\mathcal{K}-\mathcal{K})\boldsymbol{\theta}} \\ &= \mathbb{I}. \end{aligned} \tag{1.24}$$

If an operator $\boldsymbol{U}$ is unitary, it represents unitary transformations, under which a vector $|\psi\rangle$ transforms as

$$|\psi'\rangle = \boldsymbol{U}|\psi\rangle, \tag{1.25}$$

and an operator $\boldsymbol{\mathcal{A}}$ transform as

$$\boldsymbol{\mathcal{A}'} = \boldsymbol{U}\boldsymbol{\mathcal{A}}\boldsymbol{U}^{-1}. \tag{1.26}$$

The latter definition allows the vector $|\phi\rangle = \boldsymbol{\mathcal{A}}|\psi\rangle$ to transform as

$$\begin{aligned} |\phi'\rangle &= (\boldsymbol{U}\boldsymbol{\mathcal{A}}\boldsymbol{U}^{-1})\boldsymbol{U}|\psi\rangle \\ &= \boldsymbol{U}(\boldsymbol{\mathcal{A}}|\psi\rangle) = \boldsymbol{U}|\phi\rangle. \end{aligned} \tag{1.27}$$

Summarizing, for a unitary transformation $\boldsymbol{U}$,

$$\begin{aligned} \boldsymbol{U}^\dagger &= \boldsymbol{U}^{-1}, \\ |\psi'\rangle &= \boldsymbol{U}|\psi\rangle, \\ \langle\psi'| &= \langle\psi|\boldsymbol{U}^\dagger, \\ \boldsymbol{\mathcal{A}'} &= \boldsymbol{U}\boldsymbol{\mathcal{A}}\boldsymbol{U}^\dagger, \\ <\psi'|\boldsymbol{\mathcal{A}'}|\psi'\rangle &= \langle\psi|\boldsymbol{\mathcal{A}}|\psi\rangle. \end{aligned} \tag{1.28}$$

This last line emphasizes the fact that if you transform BOTH the operators and states, the matrix element is unchanged. However, in some cases you transform the states, leaving the operator unchanged, or you might transformed the operator and leave the state unchanged. This leads to new matrix elements. It can be painful to keep track of which objects are being transformed. This is akin to performing a rotation, where rotating an object by ϕ or rotating the coordinate system by $-\phi$ has the same effect. Also, if you rotate both the coordinate system and the object by the same angle, nothing changes.

1.5 Density Matrices

First, we stop and consider what forms of matrix elements might be considered as an observable. Observables must be real and independent of the basis, i.e. all unitary transformations that act

both on the bras and kets and on the operators should leave the observable unchanged. All observables can be expressed as either the expectation of a Hermitian operator,

$$\langle\psi|\mathcal{K}|\psi\rangle, \tag{1.29}$$

or as the squared overlap of two states,

$$|\langle\phi|\mathcal{U}|\psi\rangle|^2. \tag{1.30}$$

Instead of describing a state ψ by a vector, one could describe it by a density matrix,

$$\rho_\psi = |\psi\rangle\langle\psi|, \tag{1.31}$$

or as a matrix

$$(\rho_\psi)_{ij} = \psi_i\psi_j^*. \tag{1.32}$$

By inspection, one can see that the density matrix is Hermitian. From the definition of the density matrix, one can write the state $|\psi\rangle$ as a vector and the operator $\mathcal{A}$ as a matrix, and that

$$\begin{aligned}\langle\psi|\mathcal{A}|\psi\rangle &= \sum_{i,j}\psi_i^* A_{ij}\psi_j \\ &= \mathrm{Tr}\ \rho_\psi\mathcal{A},\end{aligned} \tag{1.33}$$

and that

$$\begin{aligned}|\langle\phi|\mathcal{A}|\psi\rangle|^2 &= <\psi|\mathcal{A}^\dagger|\phi\rangle\langle\phi|\mathcal{A}|\psi\rangle \\ &= \mathrm{Tr}\ \rho_\psi\mathcal{A}^\dagger\rho_\phi\mathcal{A}.\end{aligned} \tag{1.34}$$

Thus, density matrices are sufficient to generate all observables. The trace of any product of operators, or matrices, is invariant to unitary transformations so the answer should be independent of basis, as long as the density matrices and the operator $\mathcal{A}$ are all transformed.

Example 1.2: Density Matrix for a Two-Component System
Consider the state

$$|\psi\rangle = \begin{pmatrix}\cos\theta \\ e^{i\phi}\sin\theta\end{pmatrix}.$$

- Calculate the density matrix and find the trace. Then find the eigenvalues of ρ_ψ.
  Solution:
  Write down the outer product of the state,

$$\begin{aligned}\rho_\psi &= \begin{pmatrix}\cos\theta \\ e^{i\phi}\sin\theta\end{pmatrix}\begin{pmatrix}\cos\theta & e^{-i\phi}\sin\theta\end{pmatrix} \\ &= \begin{pmatrix}\cos^2\theta & e^{-i\phi}\sin\theta\cos\theta \\ e^{i\phi}\sin\theta\cos\theta & \sin^2\theta\end{pmatrix}.\end{aligned}$$

The trace is $\cos^2\theta + \sin^2\theta = 1$. The eigenvalues, λ, are found by

$$\begin{aligned} |\rho_\psi - \lambda\mathbb{I}| &= 0 \\ &= \begin{vmatrix} \cos^2\theta - \lambda & e^{-i\phi}\sin\theta\cos\theta \\ e^{i\phi}\sin\theta\cos\theta & \sin^2 - \lambda \end{vmatrix}, \\ 0 &= \lambda^2 + \cos^2\theta\sin^2\theta - \lambda(\cos^2 + \sin^2\theta) - \sin^2\theta\cos^2\theta, \\ \lambda^2 - \lambda &= 0. \end{aligned}$$

The eigenvalues are zero and 1, as expected for a pure state.

- What is ρ_ψ if one averages over all phases ϕ? Did this change the trace? What are the eigenvalues?
  Solution:
  The off-diagonal terms average to zero and

$$\rho_\psi = \begin{pmatrix} \cos^2\theta & 0 \\ 0 & \sin^2\theta \end{pmatrix}$$

The trace is still unity. The eigenvalues are $\cos^2\theta$ and $\sin^2\theta$.

For a state $|\psi\rangle$, one can always choose a basis where

$$|\psi\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{1.35}$$

In this basis ρ_ψ is diagonalized and has the form,

$$\rho_\psi = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ & & \vdots & \\ 0 & 0 & \cdots & 0 \end{pmatrix}. \tag{1.36}$$

A density matrix that can be diagonalized in such a way is known as a *pure* state. A projection operator is a density matrix corresponding to a pure state. The reason that it is often called a projection operator is that

$$P_\psi^2 = P_\psi. \tag{1.37}$$

Further, for a pure state,

$$\mathrm{Tr}\, P_\psi = 1. \tag{1.38}$$

Projection operators can also play the role of a filter, e.g. a polarization filter. If one has a state that is a linear combination of various states, $a|\psi\rangle + b|\phi\rangle$, the projection operator P_ψ acting on such a state returns $a|\psi\rangle$.

However, density matrices can be more general. If one has a density matrix which is an incoherent sum over several states, the resulting diagonalized density matrix could have more than one non-zero element, though the trace would have to remain equal to unity. Thus, density matrices can be used to express non-pure states such as unpolarized beams. For instance, one can define the following density matrix,

$$\boldsymbol{\rho} = \frac{\mathbf{1}}{\mathbf{2}}|\mathbf{1}\rangle\langle\mathbf{1}| + \frac{\mathbf{1}}{\mathbf{2}}|\mathbf{2}\rangle\langle\mathbf{2}|. \tag{1.39}$$

As with the pure state, the trace remains at unity, but after being diagonalized the strength is spread along the diagonal. This density matrix describes being in state $|\mathbf{1}\rangle$ 50% of the time and in state $|\mathbf{2}\rangle$ 50% of the time. A way to write a state that is described by this density matrix, would be

$$\frac{\mathbf{1}}{\sqrt{\mathbf{2}}}|\mathbf{1}\rangle + \boldsymbol{e^{i\phi}}\frac{\mathbf{1}}{\sqrt{\mathbf{2}}}|\mathbf{2}\rangle, \tag{1.40}$$

with ϕ being treated as a random phase to justify ignoring the off-diagonal terms in the density matrix. When performing calculations with such a state, one would ignore all terms in the density matrix with a leftover phase, $e^{i\phi}$, to account for the randomness of the phase. Note that a density matrix $\boldsymbol{\rho}$ for an impure state is not a projection operator, i.e. $\boldsymbol{\rho^2 \neq \rho}$.

Density matrices play an essential part in thermodynamics. In that case, one considers incoherent sums over many states weighted by the energy. If one is in a basis where the Hamiltonian is diagonalized, the density matrix takes the form

$$\boldsymbol{\rho_{\text{thermal}}} = \frac{\mathbf{1}}{\boldsymbol{Z}}\begin{pmatrix} \boldsymbol{e^{-\beta E_1}} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{e^{-\beta E_2}} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{e^{-\beta E_N}} \end{pmatrix}, \tag{1.41}$$

$$\boldsymbol{Z} = \sum_{\boldsymbol{i}} \boldsymbol{e^{-\beta E_i}},$$

where $\boldsymbol{\beta}$ is the inverse temperature.

One might ask the question whether it makes sense to do quantum mechanics in terms of density matrices rather than with wave functions. Density matrices are Hermitian, and given the constraint that the trace is unity, and $\boldsymbol{N \times N}$ density matrix can be represented by $\boldsymbol{N^2 - 1}$ real numbers. A wave function has $\boldsymbol{N}$ complex components, and given the normalization, have $\boldsymbol{2N - 1}$ components. Of course, there is also a meaningless phase, so a wave function can be physically represented by $\boldsymbol{2N - 2}$ numbers. Clearly, the wave function is more efficient. However, the density matrix can also account for incoherent mixtures of states. This explains the need for additional information.

1.6 Rotations of Two-Component Spin-Half Systems

As an example of a unitary transformation we consider rotations in a two-component system. For spin-1/2 systems the spin operator is

$$\vec{\boldsymbol{S}} = \frac{\hbar}{\mathbf{2}}\vec{\boldsymbol{\sigma}}. \tag{1.42}$$

For two-component systems, all operators can be written as a linear combination of the Pauli $\boldsymbol{\sigma}$ matrices and the unit matrix.

$$\boldsymbol{\sigma_z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \ \boldsymbol{\sigma_x} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \boldsymbol{\sigma_y} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{1.43}$$

As we will see later, under rotations the three Pauli matrices transform like components of a vector. For now, we only notice that each matrix is Hermitian and that when each sigma matrix is squared it gives the unit matrix. Note that the Pauli matrices have the properties,

$$\begin{aligned} \boldsymbol{\sigma_i^2} &= \mathbf{1}, \\ \boldsymbol{\{\sigma_i, \sigma_j\}} &= \boldsymbol{2\delta_{ij}}, \\ \boldsymbol{[\sigma_i, \sigma_j]} &= \boldsymbol{2i\epsilon_{ijk}\sigma_k}, \\ \boldsymbol{\sigma_i\sigma_j} &= \boldsymbol{\delta_{ij} + i\epsilon_{ijk}\sigma_k}. \end{aligned} \tag{1.44}$$

where the anti-commutator is noted by $\boldsymbol{\{A, B\} \equiv AB + BA}$ for operators $\boldsymbol{A}$ and $\boldsymbol{B}$, and the regular commutator is noted by $\boldsymbol{[A, B] \equiv AB - BA}$. From these expressions one can readily show that for a unit vector $\hat{\boldsymbol{n}}$,

$$\boldsymbol{(\vec{\sigma} \cdot \hat{n})^2 = \mathbb{I}}. \tag{1.45}$$

For $\boldsymbol{i \neq j}$ one can see

$$\begin{aligned} \boldsymbol{\sigma_i\sigma_{j\neq i}} &= \boldsymbol{i\epsilon_{ijk}\sigma_k}, \\ \boldsymbol{\sigma_x\sigma_y} &= \boldsymbol{i\sigma_z}, \ \boldsymbol{\sigma_y\sigma_z = i\sigma_x}, \ \boldsymbol{\sigma_z\sigma_x = i\sigma_y}. \end{aligned} \tag{1.46}$$

For a state with spin-up or spin-down along the $\boldsymbol{z}$ axis, we choose the basis

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{1.47}$$

Rotations by an angle $\boldsymbol{\theta}$ are given by:

$$\boldsymbol{\mathcal{R}(\vec{\theta}) = e^{-i\vec{\theta}\cdot\vec{\sigma}/2}}, \tag{1.48}$$

where the direction of $\vec{\boldsymbol{\theta}}$ is along the axis of rotation $\hat{\boldsymbol{n}}$, or equivalently, $\boldsymbol{\vec{\theta} = \theta\hat{n}}$.
From performing a Taylor expansion, and using Eq. (1.44),

$$\boldsymbol{e^{-i\theta\vec{\sigma}\cdot\hat{n}/2} = \cos(\theta/2) - i\sin(\theta/2)\vec{\sigma}\cdot\hat{n}}. \tag{1.49}$$

This trick comes in handy for a large number of physics examples, not just rotations.

Example 1.3: Rotating Spin-1/2 Systems
Consider a fermion's spin to originally be in the spin-up state, where "up" is defined by the $\boldsymbol{z}$ axis.

$$|\psi\rangle = |\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

a) Find $|\psi\rangle$ after rotating by angles of $\theta = 90,^\circ, 180^\circ$ and 360° about the x axis.

b) What are the expectations of $\langle\psi|\sigma_i|\psi\rangle$ for each rotation.

c) What is $|\psi\rangle$ if it is rotated by an angle ϕ about the z axis?

Solution:
a) Inserting σ_x into Eq. (1.49),

$$\mathcal{R}(\theta) = \cos(\theta/2) - i\sin(\theta/2)\sigma_x,$$

and the answers are:

$$\begin{aligned}\mathcal{R}(\theta = 90^\circ)|\uparrow\rangle = |\psi(90^\circ)\rangle &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}, \\ |\psi(180^\circ)\rangle &= \begin{pmatrix} 0 \\ -i \end{pmatrix}, \\ |\psi(360^\circ)\rangle &= \begin{pmatrix} -1 \\ 0 \end{pmatrix}.\end{aligned}$$

b) Just stick in the Pauli matrices,

$$\begin{aligned}\langle\psi(0^\circ)|\sigma_x|\psi(0^\circ)\rangle = 0,\ \langle\psi(0^\circ)|\sigma_y|\psi(0^\circ)\rangle = 0, \langle\psi(0^\circ)|\sigma_z|\psi(0^\circ)\rangle = 1, \\ \langle\psi(90^\circ)|\sigma_x|\psi(90^\circ)\rangle = 0,\ \langle\psi(90^\circ)|\sigma_y|\psi(90^\circ)\rangle = \frac{1}{2}, \langle\psi(90^\circ)|\sigma_z|\psi(90^\circ)\rangle = \frac{1}{2}, \\ \langle\psi(180^\circ)|\sigma_x|\psi(180^\circ)\rangle = 0,\ \langle\psi(180^\circ)|\sigma_y|\psi(180^\circ)\rangle = 0, \langle\psi(180^\circ)|\sigma_z|\psi(180^\circ)\rangle = -1, \\ \langle\psi(360^\circ)|\sigma_x|\psi(360^\circ)\rangle = 0,\ \langle\psi(360^\circ)|\sigma_y|\psi(360^\circ)\rangle = 0, \langle\psi(360^\circ)|\sigma_z|\psi(360^\circ)\rangle = 1.\end{aligned}$$

c) Just reading it off from Eq. (1.49),

$$|\psi(\phi = 90^\circ)\rangle = \begin{pmatrix} e^{-i\phi/2} \\ 0 \end{pmatrix},$$

which is simply the original state multiplied by a phase factor. One odd feature here is that the rotating by 360° returns the same state multiplied by -1.

It is important to contrast the behavior for rotating spin $1/2$ systems, which are two-component systems in a three dimensional world, to rotating normal vectors in a two-dimensional world. For normal vectors, the rotation matrix would be

$$\begin{aligned}\mathcal{R}(\phi) &= \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \\ &= e^{-i\phi\sigma_y} \\ &= \cos\phi - i\sigma_y\sin\phi.\end{aligned} \tag{1.50}$$

Note the lack of the factor $1/2$ in the argument, and that σ_y is being used even though this is a rotation about the z axis. Also, for a rotation of 360°, one returns to the same state, without the

additional factor of -1. The fact that both kinds of transformations represent the same rotations is something that will be discussed during the brief discussion of group theory in 4.2.

1.7 Polarization of Photons

Electromagnetic waves are comprised of quantized photons. For an electromagnetic wave traveling in some direction, the light can be polarized, with the electric field oscillating being in some direction perpendicular to the direction of propagation. If the light is moving in the z direction, the field might be oscillating in the x direction. For a photon the spin is quantized into two possible polarization states. For a photon propagating in the z direction, those states might be $|x\rangle$ or $|y\rangle$, which would correspond to light polarized in the x or the y directions. The two states might be represented by the vectors

$$|x\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |y\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{1.51}$$

Note that one switches from the $|x\rangle$ to the $y\rangle$ if rotated by $90°$. This is in contrast to spin-half particles, where one rotates by $180°$ to switch to the orthonormal state. Of course, the other difference is that for photons we are considering only rotations about the direction of propagation, whereas for spin-half particles, one must consider three dimensions of rotations.

The wave function for photons moving in the z direction has the form

$$\phi(x,t) = e^{-i\omega t + ikz} \begin{pmatrix} a \\ b \end{pmatrix}, \tag{1.52}$$

where the vector is known as the polarization vector, which we consider to be normalized $a^2 + b^2 = 1$. The following polarization vectors describe linearly polarized photons,

$$\begin{aligned} &\begin{pmatrix} 1 \\ 0 \end{pmatrix}, && \textbf{linear polarization along the x axis} \\ &\begin{pmatrix} 0 \\ 1 \end{pmatrix}, && \textbf{linear polarization along the y axis} \\ \tfrac{1}{\sqrt{2}} &\begin{pmatrix} 1 \\ 1 \end{pmatrix}, && \textbf{linear polarization along a direction } 45° \textbf{ from } x \textbf{ axis.} \end{aligned} \tag{1.53}$$

Another linear combination of $|x\rangle$ and $y\rangle$ states are right- and left-circularly polarized photons. Those states are

$$\begin{aligned} |R\rangle &= \frac{1}{\sqrt{2}}\left(|x\rangle + i|y\rangle\right) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}, \\ |L\rangle &= \frac{1}{\sqrt{2}}\left(|x\rangle - |y\rangle\right) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}, \end{aligned} \tag{1.54}$$

To see why these are called linear polarized photons, one uses the fact that the electric fields are proportional to the corresponding components of the wave function. Looking at those points

with $z = 0$, for the polarization with upper and lower components a and b, the electric fields behave as

$$\begin{aligned} E_x &= \mathrm{Re} E_0 a e^{-i\omega t}, \\ E_y &= \mathrm{Re} E_0 b e^{-i\omega t}. \end{aligned} \tag{1.55}$$

For $a = 1/\sqrt{2}$ and $b = i\sqrt{2}$, this becomes

$$E_x = \frac{E_0}{\sqrt{2}} \cos \omega t, \;\; E_y = \frac{E_0}{\sqrt{2}} \sin \omega t. \tag{1.56}$$

At $t = 0$ the electric field is pointed along the positive z axis, but after one fourth of a period it is pointed along the positive y axis. At this fixed value of z, the direction of the field is then rotating about the z axis and is called right circularly polarized. If one had considered the polarization vector $a = 1/\sqrt{2}$, $b = -i/\sqrt{2}$, the polarization vector would have rotated in the opposite direction, and be called left-circularly polarized. As long as one fixes the momentum of the photon, there are two polarizations and the polarization states can be described in a two-component basis. If one wants the polarization vector to describe the direction of the polarization even when the photon is propagating away from the z axis, polarization vectors must then become three-component vectors. But because polarizations are perpendicular to the direction of propagation, $\hat{p}$, one must project away the contribution of those vectors in the $\hat{p}$ direction.

Example 1.4: Photon Polarization
Photons are traveling along the z axis and are polarized along the x axis. They pass through a polarization filter which only permits passage of photons whose polarization is at an angle 30° to the x axis. What fraction of photons pass through the filter.
Solution:
Consider a single photon. The original polarization vector is

$$|\phi = 0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

and the state that passes through the filter is

$$|\phi = 30^\circ\rangle = \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}.$$

The probability of being in the new state is the squared overlap,

$$|\langle \phi = 30^\circ | \phi = 0 \rangle|^2 = \frac{3}{4}.$$

1.8 Evolution in Time

Any function of time can be expressed as a Taylor series. For a state,

$$|\psi(t_0+\tau)\rangle = |\psi(t_0)\rangle + \tau\frac{d}{dt}|\psi(t)\rangle_{t=t_0} + \frac{\tau^2}{2}\frac{d^2}{dt^2}|\psi(t)\rangle_{t=0} + \cdots + \frac{\tau^n}{n!}\frac{d^n}{dt^n}|\psi(t)\rangle_{t=t_0} + \cdots . \tag{1.57}$$

Using the fact that the Taylor expansion for $e^x = 1 + x + \cdots + x^n/n! + \cdots$, one can see that

$$\begin{aligned} |\psi(t_0+\tau)\rangle &= U(t_0, t_0+\tau)|\psi(t)\rangle_{t=t_0} \\ U(t_0, t_0+\tau) &= \exp\left\{\tau\frac{d}{dt}\right\}. \end{aligned} \tag{1.58}$$

The operation $U(t_0, t_0+\tau)$ is referred to as the evolution operator as it evolves the state forward (or backward) in time by an amount τ. The norm of a state, $\langle\psi|\psi\rangle$, represents the probability so it which must remain equal to unity. This puts a condition on U,

$$\begin{aligned} \langle\psi(t)|\psi(t)\rangle &= \langle\psi(0)|\psi(0)\rangle \\ &= \langle\psi(0)|U^\dagger(0,t)U(0,t)|\psi(0)\rangle, \\ U^\dagger(0,t)U(0,t) &= \mathbb{I}. \end{aligned} \tag{1.59}$$

Thus, the evolution operator must be a unitary operator.

For small times,

$$U(t, t+\delta t) = 1 + \delta t\frac{d}{dt}, \tag{1.60}$$

where d/dt returns the derivative w.r.t. at time t. We can define an operator H as

$$\begin{aligned} H(t) &= i\hbar\frac{d}{dt}, \\ H(t)|\psi(t)\rangle &= i\hbar\frac{d}{dt}|\psi(t)\rangle, \end{aligned} \tag{1.61}$$

which is Schrödinger's equation in general form and is nothing more than a definition of H. The evolution operator for small times is

$$\begin{aligned} U(t, t+\delta t) &\approx 1 - i\frac{H(t)}{\hbar}\delta t, \\ U^\dagger(t, t+\delta t)U(t+\delta t) &\approx 1 - i\frac{H(t) - H^\dagger(t)}{\hbar}\delta t. \end{aligned} \tag{1.62}$$

Because U is unitary, one can state that $H(t) = H^\dagger(t)$, or that $H(t)$ must be Hermitian.

If H is independent of time, one can take inspect the Taylor expansion in Eq. (1.57) and see that

$$U(t, t+\tau) = e^{-iH\tau/\hbar}. \tag{1.63}$$

For many considerations, e.g. in time-dependent perturbation theory, H depends explicitly on time, which then invalidates the simple expression above.

1.9 Evolution of Two-component Systems

Many problems in quantum mechanics can be reduced to two-state problems. Aside from the "spin-up spin-down" problem, the two-kaon problem (see Baym), the solar neutrino problem along with many other examples are really simple variations of the two-state problem.

The two-state problem is especially nice because all two-by-two matrices can be written as a linear combination of the unit matrix and the three sigma matrices. For convenience, we repeat the properties of the Pauli matrices,

$$\boldsymbol{\sigma_z} = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{-1} \end{pmatrix}, \ \boldsymbol{\sigma_x} = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix}, \ \boldsymbol{\sigma_y} = \begin{pmatrix} \mathbf{0} & \mathbf{-i} \\ \boldsymbol{i} & \mathbf{0} \end{pmatrix}, \ \mathbb{I} = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix}. \tag{1.64}$$

As presented earlier, the Pauli matrices are Hermitian, traceless, and obey simple commutation relations,

$$[\boldsymbol{\sigma_i}, \boldsymbol{\sigma_j}] = \mathbf{2}\boldsymbol{i}\boldsymbol{\epsilon_{ijk}}\boldsymbol{\sigma_k}, \ \ \{\boldsymbol{\sigma_i}, \boldsymbol{\sigma_j}\} = \mathbf{2}\boldsymbol{\delta_{ij}}. \tag{1.65}$$

In fact, as we will see later, the matrices $\hbar\vec{\boldsymbol{\sigma}}/\mathbf{2}$ obey the same commutation rules as angular momentum. From the anti-commutation relations, the square of any $\boldsymbol{\sigma}$ matrix is unity,

$$(\vec{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{n}})^{\mathbf{2}} = \mathbb{I} \tag{1.66}$$

The evolution of states under a Hamiltonian,

$$\boldsymbol{H} = \boldsymbol{\beta}\vec{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{n}}, \tag{1.67}$$

is especially simple. In this case the evolution operator is

$$\begin{aligned} \boldsymbol{U(t)} &= \boldsymbol{e^{-iHt/\hbar}} \\ &= \mathbf{1} - \boldsymbol{i\beta t}\vec{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{n}} + (\boldsymbol{-i\beta t}\vec{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{n}})^{\mathbf{2}}/\mathbf{2!} + \cdots + (\boldsymbol{i\beta t}\vec{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{n}})^{\boldsymbol{N}}/\boldsymbol{N}! + \cdots \\ &= \mathbf{cos}(\boldsymbol{\beta t}) - \boldsymbol{i}\vec{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{n}}\,\mathbf{sin}(\boldsymbol{\beta t}), \end{aligned} \tag{1.68}$$

because

$$(\vec{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{n}})^{\boldsymbol{N}} = \begin{cases} \mathbb{I}, & \boldsymbol{N} = \mathbf{even} \\ \vec{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{n}}, & \boldsymbol{N} = \mathbf{odd} \end{cases} \tag{1.69}$$

Example 1.5: Evolving Spin 1/2 Particles
A spin-up(along the z-axis) particle is placed in an environment at $\boldsymbol{t} = \mathbf{0}$ where it interacts with a magnetic field pointed along the $\boldsymbol{x}$ axis,

$$\boldsymbol{H} = \boldsymbol{\beta}\boldsymbol{\sigma_x}.$$

Find the probability of being in the "up" state as a function of time.
Solution:

$$\begin{aligned}|\psi(t)\rangle &= e^{-iHt/\hbar}|\psi(0)\rangle \\ &= (\cos(\beta t) - i\sin(\beta t)\sigma_x)\begin{pmatrix}1\\0\end{pmatrix} \\ &= \begin{pmatrix}\cos(\beta t)\\ -i\sin(\beta t)\end{pmatrix}.\end{aligned}$$

Thus, the probability of being in the "up" state is $\cos^2(\beta t)$.

Example 1.6: Neutrino Mixing
Two species of neutrinos, the ν_μ and ν_τ, have masses m_μ and m_τ. The Hamiltonian that describes these masses could be written:

$$H_0 = \begin{pmatrix} m_\mu & 0 \\ 0 & m_\tau \end{pmatrix}.$$

Now, we consider an extra term added to the Hamiltonian that mixes the two flavors of neutrinos,

$$H_{\text{mix}} = \begin{pmatrix} 0 & \alpha \\ \alpha & 0 \end{pmatrix}.$$

First, let us find the energies of two new states, which requires finding the eigenvalues of the Hamiltonian. We do this by writing the Hamiltonian in terms of the sigma matrices.

$$H = \frac{1}{2}(m_\mu + m_\tau)\mathbb{I} + \frac{1}{2}(m_\mu - m_\tau)\sigma_z + \alpha\sigma_x$$

For a three dimensional vector, one can rewrite a vector,

$$\begin{aligned}\vec{A} &= A_x\hat{x} + A_y\hat{y} + A_z\hat{z} \\ &= |\vec{A}|\hat{n}, \\ |\vec{A}| &= \sqrt{A_x^2 + A_y^2 + A_z^2}, \\ \hat{n} &= (A_x\hat{x} + A_y\hat{y} + A_z\hat{z})/|\vec{A}|.\end{aligned}$$

Because the σ matrices rotate amongst one another in the same way as $\hat{x}$, $\hat{y}$ and $\hat{z}$, we are motivated to rewrite the Hamiltonian as

$$H = \frac{1}{2}(m_\mu + m_\tau)\mathbb{I} + \beta\vec{\sigma}\cdot\hat{n}.$$

Here, β is the magnitude of the two terms that multiply sigma matrices,

$$\beta = \sqrt{\alpha^2 + (m_\tau - m_\mu)^2/4},$$

and $\hat{n}$ is a unit vector pointing in the direction,

$$\begin{aligned}\hat{n} &= \hat{z}\cos\theta + \hat{x}\sin\theta \\ \sin\theta &= \frac{\alpha}{\beta}\end{aligned}$$

Finding the eigenvalues is simply a matter of rotating the $\boldsymbol{\sigma}$ matrices so that $\hat{\boldsymbol{n}}$ is in the $\boldsymbol{z}$ direction, at which point $\vec{\boldsymbol{\sigma}} \cdot \hat{\vec{\boldsymbol{n}}} = \boldsymbol{\sigma_z}$ and the Hamiltonian becomes diagonal. The two energies are then,

$$\boldsymbol{E_\pm = \frac{1}{2}(m_\mu + m_\tau)c^2 \mathbb{I} \pm \beta c^2.}$$

One could also just diagonalize the two-by-two matrix, which would lead to the same answer.

1.10 Heisenberg and Schrödinger representations

Usually, one wishes to calculate expectations of operators, e.g. $\langle\phi|\boldsymbol{AB\cdots C}|\psi\rangle$, where the states evolve as a function of time, but the states are considered independent of time. Considering an evolution operator, $\boldsymbol{U = e^{-iHt/\hbar}}$, one can express the time development of the expectation of $\boldsymbol{AB\cdots C}$ in either of two equivalent representations,

$$\begin{aligned}\langle\phi(t)|\boldsymbol{AB\cdots C}|\psi(t)\rangle &= \langle\phi(0)|\boldsymbol{U^\dagger AB\cdots CU}|\psi(0)\rangle \\ &= \langle\phi(0)|\,\boldsymbol{U^\dagger AU\; U^\dagger BU\; \cdots U^\dagger CU}\,|\psi(0)\rangle\boldsymbol{.}\end{aligned} \tag{1.70}$$

The upper line of Eq. (1.70) is known as the Schrödinger representation, with the states evolving while the operators are fixed. The operators above will be labeled $\boldsymbol{A_S(t) = A}$, with the subscript referring to the fact that one is in the Schrödinger representation and distinguish it from the operators in Heisenberg representation. The time dependence of $\boldsymbol{A_S(t)}$ only exists if there is an explicit time dependence added to the problem, e.g. a time-dependent potential. In the lower line of Eq. (1.70), one can consider the states as being fixed, then consider the evolution of the operators, $\boldsymbol{A_H(t) = U^\dagger(t) A_S U(t)}$. Even if $\boldsymbol{A_S}$ has no time dependence, the operators $\boldsymbol{A_H(t)}$ will depend on time unless $\boldsymbol{A_S}$ commutes with the Hamiltonian.

In the Heisenberg representation, the time development of the operators can be expressed as a differential equation, where the rate of change of the operator is given by the commutation of the Hamiltonian with $\boldsymbol{A_S}$.

$$\begin{aligned}\boldsymbol{A_H(t)} &\equiv \boldsymbol{U^\dagger(t) A_S U(t)} \\ \boldsymbol{\frac{d}{dt}A_H(t)} &= \boldsymbol{\frac{\partial}{\partial t}\left[U^\dagger(t) A_S(t) U(t)\right]} \\ &= \boldsymbol{\frac{i}{\hbar}U^\dagger(t)[H(t), A_S(t)]U(t) + U^\dagger(t)\left(\frac{\partial}{\partial t}A_S(t)\right)U(t).}\end{aligned} \tag{1.71}$$

Here, the subscripts $\boldsymbol{S}$ and $\boldsymbol{H}$ refer to Schrödinger and Heisenberg representations respectfully. If there is no explicit time dependence in $\boldsymbol{A_S}$, then any operator that commutes with the Hamiltonian represents a constant of the motion. The most obvious such operator is $\boldsymbol{H}$ itself. Thus, if the Hamiltonian has no explicit time dependence, the expectation of $\boldsymbol{H}$, a.k.a. the energy, is a constant of the motion.

Example 1.7: Spin Precession in Heisenberg Representation
The spin-precession example from before can also be described by considering how the opera-

tors, in the Heisenberg representation, change with time, as opposed to how the states change with time in the Schrödinger representation. Again, we consider the Hamiltonian, $\boldsymbol{H = \beta\sigma_x}$. The time development of $\boldsymbol{\sigma_z(t)}$ and $\boldsymbol{\sigma_y(t)}$ in the Heisenberg representation could be written as,

$$\begin{aligned}\frac{\boldsymbol{d}}{\boldsymbol{dt}}\boldsymbol{\sigma_z(t)} &= \frac{\boldsymbol{i}}{\hbar}\boldsymbol{U^\dagger(t)[H,\sigma_z]U(t)} \\ &= \frac{\boldsymbol{i}}{\hbar}\boldsymbol{\beta U^\dagger(t)[\sigma_x,\sigma_z]U(t)} = \frac{\boldsymbol{2\beta}}{\hbar}\boldsymbol{\sigma_y(t)} \\ \frac{\boldsymbol{d}}{\boldsymbol{dt}}\boldsymbol{\sigma_y(t)} &= \frac{\boldsymbol{i}}{\hbar}\boldsymbol{\beta U^\dagger(t)[\sigma_x,\sigma_y]U(t)} = -\frac{\boldsymbol{2\beta}}{\hbar}\boldsymbol{\sigma_z(t)},\end{aligned} \tag{1.72}$$

or equivalently,

$$\frac{\boldsymbol{d^2}}{\boldsymbol{dt^2}}\boldsymbol{\sigma_z(t)} = -\frac{\boldsymbol{4\beta^2}}{\boldsymbol{\hbar^2}}\boldsymbol{\sigma_z(t).} \tag{1.73}$$

Thus, the expression for $\langle\boldsymbol{\sigma_z(t)}\rangle$ must have the form

$$\langle\boldsymbol{\sigma_z(t)}\rangle = \boldsymbol{A}\cos\left(\boldsymbol{2\beta t/\hbar}\right) + \boldsymbol{B}\sin\left(\boldsymbol{2\beta t/\hbar}\right), \tag{1.74}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are chosen to satisfy the initial conditions. If the initial spin is along the $\boldsymbol{z}$ axis $\boldsymbol{A = 1}$ and $\boldsymbol{B = 0}$.

Finally, we should point out that for any two Hermitian operators $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$ that commute, a set of states can be found that are eigenstates of both $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$. To see this, consider eigenstates of $\boldsymbol{\mathcal{B}}$,

$$\boldsymbol{\mathcal{B}|b\rangle = b|b\rangle.} \tag{1.75}$$

If $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$ commute,

$$\boldsymbol{\langle b'|\left[\mathcal{A},\mathcal{B}\right]|b\rangle = 0 = (b - b')\langle b'|\mathcal{A}|b\rangle.} \tag{1.76}$$

Thus, the operator $\boldsymbol{\mathcal{A}}$ does not mix states with different eigenvalues of $\boldsymbol{\mathcal{B}}$. Or in other words, one can take the subset of states which have a given eigenvalue $\boldsymbol{b}$, and this set should provide a basis for simultaneously diagonalizing $\boldsymbol{\mathcal{A}}$.

The implications of this simple statement are profound. For example if the angular momentum operator $\boldsymbol{\mathcal{L}_z}$ commutes with the Hamiltonian, the subset of states with the same eigenvalue of $\boldsymbol{\mathcal{L}_z}$, usually denoted by $\boldsymbol{m}$, can be considered by themselves while diagonalizing the Hamiltonian.

This also implies that for any operator $\boldsymbol{\mathcal{A}}$ and a Hamiltonian $\boldsymbol{H}$, where there is no explicit time dependence in either operator, or equivalently the operator has no time dependence in the Schrödinger representation, the time rate of change of $\boldsymbol{\mathcal{A}_H(t)}$ is,

$$\begin{aligned}\frac{\boldsymbol{d}}{\boldsymbol{dt}}\boldsymbol{\mathcal{A}_H(t)} &= \frac{\boldsymbol{d}}{\boldsymbol{dt}}\left(\boldsymbol{e^{iHt/\hbar}\mathcal{A}_S e^{-iHt/\hbar}}\right) \\ &= \frac{\boldsymbol{i}}{\hbar}\boldsymbol{e^{iHt}([H,\mathcal{A}_s])e^{-iHt}.}\end{aligned} \tag{1.77}$$

Thus, any operator that commutes with the Hamiltonian is a constant of the motion.

1.11 Uncertainty Relations

In many examples one is interested in the expectations of the product of two or more Hermitian operators. For instance, one might be interested in both the Hamiltonian and the number operator. Here, we consider two Hermitian operators, $\mathcal{A}$ and $\mathcal{B}$. One can choose states to be eigenstates of either operator, but not necessarily both. This depends on whether the two operators commute. If there existed a basis where all the basis states were eigenstates of both $\mathcal{A}$ and $\mathcal{B}$, one could state that

$$\begin{aligned} \mathcal{A}\mathcal{B}|ab\rangle &= ab|ab\rangle, \\ \mathcal{B}\mathcal{A}|ab\rangle &= ab|ab\rangle. \end{aligned} \tag{1.78}$$

In this case the products $\mathcal{A}\mathcal{B}$ and $\mathcal{B}\mathcal{A}$ would return the same result acting on any basis state and one could say

$$\begin{aligned} \mathcal{A}\mathcal{B} &= \mathcal{B}\mathcal{A}, \\ [\mathcal{A},\mathcal{B}] &= 0. \end{aligned} \tag{1.79}$$

Thus, the commutation of operators is essential if one is to be able to find a basis that serves as eigenstates for both operators.

For the case where the operators no longer commute, one can find an uncertainty relation. The non-commutation results in some non-zero operator $\mathcal{C}$,

$$[\mathcal{A},\mathcal{B}] = i\mathcal{C}. \tag{1.80}$$

The operator $\mathcal{C}$ must be Hermitian because the commutation of two Hermitian operators must be anti-Hermitian,

$$\begin{aligned} (\mathcal{A}\mathcal{B}-\mathcal{B}\mathcal{A})^\dagger &= (\mathcal{B}\mathcal{A}-\mathcal{A}\mathcal{B}) \\ &= -(\mathcal{A}\mathcal{B}-\mathcal{B}\mathcal{A}). \end{aligned} \tag{1.81}$$

Equivalently, multiplying the commutation by i results in a Hermitian operator.

To derive the uncertainty relation, one begins with the Cauchy-Schwarz inequality, https://en.wikipedia.org/wiki/Cauchy\T1\textendashSchwarz_inequality. For any two operators $\delta\mathcal{A}$ and $\delta\mathcal{B}$ sandwiched between any state $|\psi\rangle$,

$$|\langle\psi|\delta\mathcal{A}\delta\mathcal{B}|\psi\rangle|^2 \leq \langle\psi|\delta\mathcal{A}^2|\psi\rangle\langle\psi|\delta\mathcal{B}^2|\psi\rangle. \tag{1.82}$$

This is related to the fact that for vectors $\vec{P}$ and $\vec{Q}$, $|\vec{P}|^2|\vec{Q}|^2 \geq (\vec{P}\cdot\vec{Q})^2$. For our purposes, we define

$$\delta\mathcal{A} \equiv \mathcal{A} - \langle\psi|A|\psi\rangle, \quad \delta\mathcal{B} \equiv \mathcal{A} - \langle\psi|B|\psi\rangle. \tag{1.83}$$

Next, one can see that

$$\begin{aligned} \langle\psi|\delta\mathcal{A}\delta\mathcal{B}|\psi\rangle &= \frac{1}{2}\langle\psi|[\mathcal{A},\mathcal{B}]|\psi\rangle + \frac{1}{2}\langle\psi|\{\delta\mathcal{A},\delta\mathcal{B}\}|\psi\rangle, \\ &= \frac{i}{2}\langle\psi|\mathcal{C}|\psi\rangle + \frac{1}{2}\langle\psi|\{\delta\mathcal{A},\delta\mathcal{B}\}|\psi\rangle. \end{aligned} \tag{1.84}$$

The first term is imaginary and the second term is real because the anti-commutator of two Hermitian operators is Hermitian. Thus,

$$|\langle\psi|\delta\mathcal{A}\delta\mathcal{B}|\psi\rangle|^2 = \frac{1}{4}\langle\psi|\mathcal{C}|\psi\rangle^2 + \frac{1}{4}\langle\psi|\{\delta\mathcal{A},\delta\mathcal{B}\}|\psi\rangle^2. \tag{1.85}$$

Plugging this into the Cauchy-Schwarz inequality, one obtains the Schrödinger-Robertson uncertainty relation,

$$\langle\psi|\delta\mathcal{A}^2|\psi\rangle\langle\psi|\delta\mathcal{B}^2|\psi\rangle \geq \frac{1}{4}\langle\psi|\mathcal{C}|\psi\rangle^2 + \frac{1}{4}\langle\psi|\{\delta\mathcal{A},\delta\mathcal{B}\}|\psi\rangle^2. \tag{1.86}$$

Both terms on the r.h.s. are manifestly positive. The expectations $\langle\psi|\delta\mathcal{A}^2|\psi\rangle$ and $\langle\psi|\delta\mathcal{B}^2|\psi\rangle$ represent the degree to which either operator is well determined by $|\psi\rangle$. If the state $|\psi\rangle$ were an eigenstate of $\mathcal{A}$, $\delta\mathcal{A}$ would be zero and the l.h.s. would vanish. Because the r.h.s. is positive, this would require $\langle\psi|\delta\mathcal{B}^2|\psi\rangle$ to be infinite. Thus, one cannot simultaneously find any state where both $\mathcal{A}$ and $\mathcal{B}$ are arbitrarily well determined.

The usual uncertainty relation involves ignoring the latter term, but including the latter term strengthens the relation. For the famous Heisenberg uncertainty relation, the operators are the position and momentum operators, $\mathcal{X}$ and $\mathcal{P} = -i\hbar\partial_x$, so $\mathcal{C} = \hbar$, and ignoring the last term in the Schrödinger-Robertson uncertainty relation,

$$\langle\psi|\delta X^2|\psi\rangle\langle\psi|\delta P^2|\psi\rangle \geq \frac{\hbar^2}{4}. \tag{1.87}$$

Example 1.8: Uncertainty Relations for Angular Momentum
Find the uncertainty relation for the angular momenta S_x and S_y.
Solution: Use the fact that $[S_x, S_y] = i\hbar S_z$ (which is true for any spin, not just spin 1/2). Here,

$$\mathcal{C} = -i[S_x, S_y] = \hbar S_z.$$

Further, $\{S_x, S_y\} = 0$. The uncertainty relation is thus

$$\langle\psi|\delta S_x^2|\psi\rangle\langle\psi|\delta S_y^2|\psi\rangle \geq \frac{\hbar^2}{4}\langle\psi|\delta S_z^2|\psi\rangle.$$

1.12 Exercises

1. Photons, traveling along the z axis can be polarized either linearly along the x or y axis, or a linear combination of the two states.

   (a) Write the operator that rotate states by 45° about the z axis in terms of $|x\rangle$, $|y\rangle$ and the corresponding bras.

   (b) Continuing with the photon moving in the z direction, and choosing the basis,

   $$|x\rangle \to \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |y\rangle \to \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

   write the matrix that rotates the states by ϕ about the z axis.

(c) Right-hand circularly polarized (RCP) light is made of a linear combination of x and y polarized light.

$$|R\rangle = \frac{1}{\sqrt{2}}\left(|x\rangle + i|y\rangle\right).$$

Light traveling along the z axis passes through a thin slab of thickness Z whose index of refraction, $k = n\omega/c$, is different for light polarized in the x and y directions. In terms of n_x, n_y and Z find the polarization vector for light which enters the slab as right-circularly polarized.
HINT: The wave has a form $e^{-i\omega t+ikz}$. The two components have the same ω but different k while in the medium.

(d) Find the density matrix for right-circularly polarized light in the basis defined above.

(e) Using the basis described above, write the density matrix for light that is an incoherent mixture, 50% polarized along the x direction and 50% along the y direction.

2. Considering a photon's polarization for a photon moving in the z direction, calculate $\langle x|\mathcal{R}(\phi)|x\rangle$ for $\phi = \pi/2, \pi, 2\pi$, where the rotation is about the z axis.

3. For a spin 1/2 particle, calculate $\langle z,+|\mathcal{R}(\theta)|z,+\rangle$, for the same angles, $\theta = \pi/2, \pi, 2\pi$, when the rotation is about the y axis.

4. Show that the unit matrix $\mathbb{I}$, which can be considered as an operator, is unchanged by a unitary transformation. Begin with the fact that for any matrix $\mathcal{M}$, $\mathcal{M}\mathbb{I} = \mathbb{I}\mathcal{M} = \mathcal{M}$.

5. Consider the rotation matrix for rotating Pauli spinors by an angle $90°$ about the z axis. Using Eq. (1.49),

$$U = e^{-i\sigma_z\pi/4} = \frac{1}{\sqrt{2}}(1 - i\sigma_z).$$

(a) Using the commutator and anti-commutator relations for the σ matrices, show that the transformation of σ_x is

$$U\sigma_x U^\dagger = \sigma_y.$$

(b) Show that rotating the state, $|+,x\rangle$, which refers to an eigenstate of σ_x with eigenvalue of +1, gives

$$U|+,x\rangle = |+,y\rangle,$$

which is the eigenstate of σ_y with eigenvalue +1.

6. Consider some Hermitian $N \times N$ matrix K_{ij}, with eigenvalues $\lambda^{(n)}$ and the corresponding normalized eigenvectors $v^{(n)}$,

$$Kv^{(n)} = \lambda^{(n)}v^{(n)}.$$

The N eigenvectors each have N components, $v_i^{(n)}$. Create an $N \times N$ matrix

$$U_{ij} = v_j^{*(i)}.$$

Thus, one is making a matrix by having each row be one of the eigenvectors.

(a) Show that U is unitary.

(b) Show that the j^{th} component of the vector $\boldsymbol{U}\boldsymbol{v}^{(n)}$ is

$$(\boldsymbol{U}\boldsymbol{v}^{(n)})_j = \boldsymbol{\delta}_{nj},$$

Thus, the vectors $\boldsymbol{U}\boldsymbol{v}^{(n)}$ are

$$(\boldsymbol{U}\boldsymbol{v}^{(1)}) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad (\boldsymbol{U}\boldsymbol{v}^{(2)}) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \cdots$$

Now, consider the matrix

$$\boldsymbol{K'} = (\boldsymbol{U}\boldsymbol{K}\boldsymbol{U}^{\dagger}),$$

and have it act on the vectors above. Show that

$$\boldsymbol{K'}(\boldsymbol{U}\boldsymbol{v}^{(n)}) = \boldsymbol{\lambda}_n(\boldsymbol{U}\boldsymbol{v}^{(n)}).$$

This shows that the vectors $(\boldsymbol{U}\boldsymbol{v}^{(n)})$ are eigenvectors of the matrix $\boldsymbol{K'}$ with eigenvalues $\boldsymbol{\lambda}_n$. Given that the eigenvectors are of the simple form above, the matrix $(\boldsymbol{U}\boldsymbol{K}\boldsymbol{U}^{\dagger})$ must be diagonal. Thus, the matrix $\boldsymbol{U}$ defined above provides the unitary matrix for transforming the matrix $\boldsymbol{K}$ into its diagonal form.

7. Consider the matrix:

$$\boldsymbol{\mathcal{M}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

(a) What are the eigenvalues of $\boldsymbol{\mathcal{M}}$?

(b) Find eigenvectors of $\boldsymbol{\mathcal{M}}$?

8. Consider the 2×2 matrix

$$\boldsymbol{\mathcal{K}} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{C}^* \\ \boldsymbol{C} & \boldsymbol{B} \end{pmatrix}$$

(a) What are the eigenvalues of $\boldsymbol{\mathcal{K}}$?
(b) What are the eigenvectors of $\boldsymbol{\mathcal{K}}$?

9. A beam of light with wavelength 660 $\boldsymbol{nm}$ is sent along the z axis through a polaroid filter that passes only x polarized light. The beam is initially polarized at 30° to the x axis, and the total energy of the pulse is exactly 10 Joules. Estimate the fluctuations of the energy of the transmitted beam, $\langle(\boldsymbol{E} - \bar{\boldsymbol{E}})^2\rangle^{1/2}$. Express the fluctuations as a fraction of the average transmitted energy. (Hint: Consider the binomial distribution, with $\boldsymbol{N}$ tries with probability $\boldsymbol{p}$ of success of each try.)

10. Consider light moving along the $\boldsymbol{z}$ axis and using the following definitions for $|\boldsymbol{R}\rangle$ and $|\boldsymbol{L}\rangle$ in terms of $\boldsymbol{x}$ and $\boldsymbol{y}$ polarized light,

$$|\boldsymbol{R}\rangle \equiv \frac{1}{\sqrt{2}}(|\boldsymbol{x}\rangle + \boldsymbol{i}|\boldsymbol{y}\rangle), \ \ |\boldsymbol{L}\rangle \equiv \frac{1}{\sqrt{2}}(|\boldsymbol{x}\rangle - \boldsymbol{i}|\boldsymbol{y}\rangle),$$

(a) In terms of $|\boldsymbol{R}\rangle$(RCP) and $|\boldsymbol{L}\rangle$(LCP) write the states $|\mathbf{45}\rangle$ and $|\mathbf{135}\rangle$ which are linearly polarized at 45° and 135° relative to the $\boldsymbol{x}$ axis.

(b) Calculate the $\mathbf{2}\times\mathbf{2}$ transformation matrix from the $\mathbf{45, 135}$ basis, where

$$|\mathbf{45}\rangle = \begin{pmatrix} \mathbf{1} \\ \mathbf{0} \end{pmatrix},\ |\mathbf{135}\rangle = \begin{pmatrix} \mathbf{0} \\ \mathbf{1} \end{pmatrix},$$

to the $\boldsymbol{R}$, $\boldsymbol{L}$ basis.

(c) Show that this transformation is unitary.

11. The probability that a photon in state $|\boldsymbol{\Psi}\rangle$ passes through an x-polaroid is the average value of a physical observable which might be called the *x-polarizedness*..

(a) Write down the operator $\boldsymbol{P_x}$, as a matrix in the $\boldsymbol{XY}$ basis where

$$|\boldsymbol{X}\rangle = \begin{pmatrix} \mathbf{1} \\ \mathbf{0} \end{pmatrix},\ |\boldsymbol{Y}\rangle = \begin{pmatrix} \mathbf{0} \\ \mathbf{1} \end{pmatrix}.$$

The projection $\langle\boldsymbol{\Psi}|\boldsymbol{P_x}|\boldsymbol{\Psi}\rangle$ is the probability that $|\boldsymbol{\Psi}\rangle$ makes it through the filter.

(b) What are its eigenvalues and eigenstates?

(c) Write the matrix $\boldsymbol{P_x}$ in the $\boldsymbol{RL}$ basis, where RCP and LCP states are

$$|\boldsymbol{R}\rangle = \begin{pmatrix} \mathbf{1} \\ \mathbf{0} \end{pmatrix},\ |\boldsymbol{L}\rangle = \begin{pmatrix} \mathbf{0} \\ \mathbf{1} \end{pmatrix},$$

and show that the eigenvalues are the same as in the $\boldsymbol{XY}$ basis. Also, show that the this matrix is a projection operator by explicitly multiplying $\boldsymbol{P_x}$ by itself.

12. The trace of a matrix $\boldsymbol{A}$ is defined as:

$$\mathbf{Tr}\boldsymbol{A} \equiv \sum_i \boldsymbol{A_{ii}}$$

(a) Show that the trace of $\boldsymbol{A}$ is invariant under a transformation of basis,

$$\boldsymbol{A} \to \boldsymbol{U^\dagger A U}$$

(b) Show that $\mathbf{Tr}\boldsymbol{AB} = \mathbf{Tr}\boldsymbol{BA}$.

13. A plane polarized photon at $\boldsymbol{\theta} = \mathbf{45}°$ enters a special crystal with indices of refraction:
$\boldsymbol{n_x}$=1.50 for photons polarized along the $\boldsymbol{x}$ axis
$\boldsymbol{n_y}$=1.52 for photons polarized along the $\boldsymbol{y}$ axis.
Assuming the wavelength of the light is 660 $\boldsymbol{nm}$ before it enters the crystal, choose the thickness of the crystal such that the outgoing light is right circularly polarized. Assume the dispersion is linear, $\boldsymbol{k = n\omega/c}$.

14. Consider the matrix for rotation about the $\boldsymbol{z}$ axis,

$$\boldsymbol{R}(\phi) = \boldsymbol{e^{-i\sigma_z\phi/2}}. \tag{1.88}$$

Show that after rotation about the $\boldsymbol{z}$ axis,

$$\boldsymbol{R}(\phi)\boldsymbol{\sigma_x}\boldsymbol{R^{-1}}(\phi) = \boldsymbol{\sigma_x}\cos(\phi) + \boldsymbol{\sigma_y}\sin(\phi) \tag{1.89}$$

15. Consider a basis for spin-up and spin-down electrons (along the z axis),

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

(a) Write down the 4 vectors describing an electron with spin pointed along the positive/negative directions of x and y axes.
(b) Write the six density matrices describing electrons polarized along the positive/negative directions of each of the three axes.
(c) Write the density matrix describing an incoherent mixture of 60% spin-up and 40% spin down.
(d) Using the density matrix, calculate $< y, +|S_z|y, +\rangle$.

16. Neutral Kaon Oscillations: There are two kinds of neutral kaons one can make using down and strange quarks,

$$|K^0\rangle = |d\bar{s}\rangle, \ |\bar{K}^0\rangle = |s\bar{d}\rangle.$$

If it weren't for the weak interaction, the two species would have equal masses, and the Hamiltonian (for a kaon with zero momentum) would be

$$H_0 = \begin{pmatrix} M & 0 \\ 0 & M \end{pmatrix}.$$

However, there is an additional term from the weak interaction that mixes the states,

$$H_m = \begin{pmatrix} 0 & \epsilon \\ \epsilon & 0 \end{pmatrix}.$$

The masses of a neutral kaon are 497.6 MeV, without mixing, but after adding the mixing term the masses differ by 3.56μeV. The two eigenstates are known as K_S (K-short) and K_L (K-long), because they decay with quite different lifetimes.

(a) What is ϵ?
(b) If one creates a kaon in the K_0 state at time $t = 0$, find the probability it would be measured as a $\bar{K}^0$ as a function of time.
(c) A beam kaons is created in the K_0 channel and has a kinetic energy of 600 MeV per kaon. Plot the probability that the kaon is in the K_0 state as a function of the distance traveled, x. Ignore the fact that the kaons decay.
(d) Repeat (c), but take into account the decays. The states

$$|K_S\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle + |\bar{K}^0\rangle,$$
$$|K_L\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle - |\bar{K}^0\rangle,$$

known at K-short and K-long, represent the eigenstates of the Hamiltonian. The lifetime of a K_L is 51.2 ns, and the lifetime of the K_S is 0.0896 ns. Note that the wave function should be modified by the factor $e^{-t/(2\tau)}$ to take decays into account decays of lifetime τ. It is often convenient to remember that $\hbar c$=197.327 MeV fm = 197.327 eV nm.

FYI: If the above were exactly true, the K_S state would be even under CP while the K_L would be odd under CP. Here, CP is an operator that changes particles to anti-particles. If the particle-antiparticle symmetry were exact, the CP operator would commute with the Hamiltonian and the eigenstates of the Hamiltonian, K_S and K_L, would have to be eigenstates of CP. The K_S would then decay to two pions and the K_L could decay to three pions. However, there is an additional small CP violating term in the Hamiltonian which allows K_L to have a small probability of decaying to two pions. This was the first experimental laboratory observation of CP violation. CP violation is required to explain the preponderance of matter vs. anti-matter in the universe.

17. Neutrino Oscillations: There are three kinds of neutrinos corresponding to the three lepton families, and recent evidence has suggested that they may oscillate between generations. Here we consider two flavors, the μ neutrino and the τ neutrino. Suppose that the Hamiltonian can be written as a free term plus a term that mixes the μ and τ neutrinos, which is proportional to α.

$$\mathcal{H} = \begin{pmatrix} \sqrt{k^2 + m_\mu^2} & 0 \\ 0 & \sqrt{k^2 + m_\tau^2} \end{pmatrix} + \alpha \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

(a) Supposing you are in the rest frame of the neutrino and that the momentum k is zero, show that the evolution operator $e^{-i\mathcal{H}t/\hbar}$ can be written as

$$e^{-i(m_\mu + m_\tau)t/2\hbar} \left\{ \cos \omega t - i \sigma_n \sin \omega t \right\},$$

where

$$\hbar\omega \equiv \sqrt{\alpha^2 + (\frac{m_\tau - m_\mu}{2})^2}$$

$$\sigma_n \equiv \frac{\frac{m_\mu - m_\tau}{2}\sigma_z + \alpha \sigma_x}{\hbar\omega}$$

(b) If a neutrino starts as a μ neutrino, what is the probability, as a function of time, of being a τ neutrino?

(c) As a function of the masses and α, what is the oscillation time? I.e. the time to return to its original flavor.

(d) If the neutrinos are extremely relativistic, $k >> m$, describe how the oscillation time translates into an oscillation as a function of the distance from the creation.

Note: Here the "masses" are the rest energies (mc^2).

18. Show that

$$\mathrm{Tr} A_S B_S C_S = \mathrm{Tr} A_H(t) B_H(t) C_H(t),$$

where the subscripts refer to Schrödinger and Heisenberg representations.

2 Coordinate and Momentum Space

In the previous section we considered problems with a finite number of discrete states. If there were N discrete states, the problem was reduced to a linear algebra exercise involving N-dimensional vectors to represent the states, and $N \times N$ dimensional matrices to represent the operators. For particles confined to a finite volume, such as a harmonic oscillator or a infinite potential well, the system still has discrete levels, but the number of such levels is unbounded as the states can have arbitrarily high energy. A vector space, where the dimensionality approaches infinity is known as a *Hilbert space*. Thus, linear algebra is not applicable without some sort of truncation. In its place, differential equations often provide the dominant mathematical tool to represent the physics.

Having the number of states approach infinity is typically the result of either (a) a continuum of points in coordinate space or (b) an infinite extent to coordinate space, or (c) both. An example of (a) is an infinite square well or a harmonic oscillator. The solutions to these problems lead to an infinite number of discrete states. An example of (b) would be an infinitely long chain with sites separated by some distance ℓ. In this case there is a continuum of momentum solutions, e.g. phonons, but with a maximum momentum of the order $\hbar/\ell$. Finally, for (c), one has a continuum of solutions, and no upper bound on the energy or momentum. For problems with a continuum in coordinate space, the physics is often described by some sort of wave equation, most famously Schrödinger's wave equation. The qualifier *wave* is invoked to distinguish it from $H\psi = E\psi$, which is not necessarily a wave equation, and might also be referred to as Schrödinger's equation.

2.1 Schrödinger's Wave Equation

One of the most famous equations in physics is Schrödinger's wave equation,

$$i\hbar\frac{\partial}{\partial t}\psi(x) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x) + V(x)\psi(x), \tag{2.1}$$

where ψ is a complex function of the continuous variable x. Rather than thinking of $\psi(x)$ as a wave function, it is more revealing to recognize it as overlap of the state $|\psi\rangle$ with the state $|x\rangle$.

$$\psi(x) = \langle x|\psi\rangle. \tag{2.2}$$

The only difference between a label in coordinate space and a label that denotes a discrete variable such as spin, is that because the x label is continuous, the normalization has to be changed.

$$\langle x'|x\rangle = \delta(x - x'), \tag{2.3}$$

where $\delta(x-x')$ is a Dirac delta function, as opposed to the usual Kronecker delta function, δ_{mn}, used for discrete states. Dirac delta functions are zero when $x \neq x'$ and infinty for $x = x'$ such that

$$\int dx'\, \delta(x - x') = 1, \tag{2.4}$$

whereas for Kronecker deltas

$$\sum_n \delta_{mn} = 1. \tag{2.5}$$

Equation (2.3) implies that $|x\rangle$ has dimensions of inverse length to the one-half power. The completeness relation becomes

$$\int dx\ |x\rangle\langle x| = \mathbb{I}, \tag{2.6}$$
$$\int dx\ \langle\phi|x\rangle\langle x|\psi\rangle = \langle\phi|\psi\rangle.$$

Next, we wish to show that the Schrödinger equation is merely the continuum limit of a matrix equation, with $\psi_0, \psi_1 \cdots \psi_N$ representing $\psi(x)$ at $x = 0, \delta x, 2\delta x \cdots$. To do this we write the second derivative as

$$\begin{aligned}\frac{\partial^2}{\partial x^2}\psi(x) &= \lim_{\delta x\to 0}\frac{\partial}{\partial x}\frac{\psi(x+\delta x/2)-\psi(x-\delta x/2)}{\delta x} \\ &= \lim_{\delta x\to 0}\frac{\psi(x+\delta x)-2\psi(x)+\psi(x-\delta x)}{\delta x^2}.\end{aligned} \tag{2.7}$$

Thus, by making the substitution,

$$\psi(x) \to \psi_i\frac{1}{\sqrt{\delta x}}, \tag{2.8}$$

the normalization becomes:

$$\sum_i |\psi_i|^2 = 1. \tag{2.9}$$

One may now rewrite Schrödinger's equation in terms of discrete vectors,

$$i\frac{\partial}{\partial t}\psi_i(t) = -\frac{\hbar^2}{2m\delta x^2}\left(\psi_{i+1}(t) - 2\psi_i(t) + \psi_{i-1}(t)\right) + V(x)\psi_i(t). \tag{2.10}$$

If one takes the limit of $\delta x \to 0$, the number of states $|x\rangle$ within a finite length approaches infinity.

For the case where δx is finite one may also write the Hamiltonian as a matrix.

$$H = -\frac{\hbar^2}{2m\delta x^2}\begin{pmatrix} \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & -2 & 1 & 0 & \cdots \\ \cdots & 1 & -2 & 1 & \cdots \\ \cdots & 0 & 1 & -2 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix} + \begin{pmatrix} \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & V(x_{i-1}) & 0 & 0 & \cdots \\ \cdots & 0 & V(x_i) & 0 & \cdots \\ \cdots & 0 & 0 & V(x_{i+1}) & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}. \tag{2.11}$$

Here, the potential term is diagonal while the kinetic term is *band diagonal*.

2.2 Momentum States

A momentum state is no more than a linear combination of coordinate-space states,

$$\begin{aligned} |p\rangle &= \int dx\; e^{ipx/\hbar}|x\rangle, \\ \langle x|p\rangle &= \int dx'\; e^{ipx'/\hbar}\langle x|x'\rangle \\ &= \int dx'\; e^{ipx'/\hbar}\delta(x-x') \\ &= e^{ipx/\hbar}. \end{aligned} \tag{2.12}$$

With this definition of the state $|p\rangle$, the normalization becomes

$$\begin{aligned} \langle p'|p\rangle &= \int dx dx'\; e^{i(px-p'x')/\hbar}\langle x'|x\rangle \\ &= \int dx\; e^{i(p-p')x/\hbar} \\ &= 2\pi\hbar\delta(p-p'), \end{aligned} \tag{2.13}$$

and the dimensionality of $|p\rangle$ is length to the one-half power.

Aside: Fourier Transforms and Delta Functions
The relation above is often used to manipulate Fourier transforms. To derive the fact that integrating the phase e^{ikx} gives a delta function, one can add a factor $e^{-\epsilon|x|}$ to the integrand and take the limit $\epsilon \to 0$. The proof then proceeds,

$$\begin{aligned} \int_{-\infty}^{\infty} dx\; e^{ikx} &= \int_{-\infty}^{0} dx\;\; e^{\epsilon x+ikx} + \int_{0}^{\infty} dx\; e^{-\epsilon x+ikx}\Big|_{\epsilon\to 0} \\ &= \frac{1}{\epsilon+ik} + \frac{1}{\epsilon-ik} \\ &= \frac{2\epsilon}{\epsilon^2+k^2}\Big|_{\epsilon\to 0}. \end{aligned} \tag{2.14}$$

The r.h.s. is clearly zero when $k \neq 0$, but is infinity for $k = 0$. To see that this is proportional to $\delta(k)$, one need only integrate the function and see,

$$\int_{-\infty}^{\infty} dk\; \frac{2\epsilon}{\epsilon^2+k^2} = 2\pi\delta(k). \tag{2.15}$$

To perform this last integral, one can do a trigonometric substitution, $k = \epsilon\tan\theta$.

The inverse transformation is

$$|x\rangle = \frac{1}{2\pi\hbar}\int dp e^{-ipx/\hbar}|p\rangle. \tag{2.16}$$

Expressed as a completeness relation,

$$\int \frac{dp}{(2\pi\hbar)}\langle x|p\rangle\langle p|x'\rangle = \int \frac{dp}{(2\pi\hbar)} e^{ip(x'-x)/\hbar} \\ = \delta(x'-x) = \langle x|x'\rangle. \tag{2.17}$$

Thus,

$$\frac{1}{2\pi\hbar}\int dp\ |p\rangle\langle p| = \mathbb{I}. \tag{2.18}$$

Sandwiching this between some normalized state ψ,

$$\frac{1}{2\pi\hbar}\int dp\ \langle\psi|p\rangle\langle p|\psi\rangle = \langle\psi|\psi\rangle \tag{2.19}$$
$$\frac{1}{2\pi\hbar}\int dp\ \psi^*(p)\psi(p) = 1,$$
$$\psi(p) = \langle p|\psi\rangle.$$

Thus, the probability per differential momentum for observing a particle with momentum p in some state ψ is

$$\frac{dP_\psi}{dp} = \frac{1}{2\pi\hbar}\psi^*(p)\psi(p). \tag{2.20}$$

Changing the problem to n dimensions only affects the expressions here by changing dp to $d^n p$ and replacing $(2\pi\hbar)$ with $(2\pi\hbar)^n$.

The label p here refers to a continuum of momentum states. If a system is confined, then no eigenstate of the momentum operator really exists, but one still uses the label p to point to eigenstates of the momentum operator within the volume, but with discrete values of p rather than a continuum. In that case the normalizations are different. It is the duty of the watchful reader to accurately interpret the notation.

Momentum and position are also operators, and as operators can be expressed in terms of bras and kets.

$$\mathcal{X} = \int dx\ x|x\rangle\langle x| \tag{2.21}$$
$$\mathcal{P} = \int \frac{dp}{2\pi\hbar}\ p|p\rangle\langle p|.$$

From the definition of $\mathcal{X}$, one can see that

$$\langle\phi|\mathcal{X}|\psi\rangle = \int dx\ \langle\phi|x\rangle x\langle x|\psi(x)\rangle \tag{2.22}$$
$$= \int dx\ \phi^*(x)x\psi(x).$$

One can also perform a similar operation with the momentum operator.

$$\begin{aligned} \langle\phi|\mathcal{P}|\psi\rangle &= \int \frac{dp}{2\pi\hbar}\phi^*(p)p\psi(p), \\ \psi(p) &= \langle p|\psi\rangle. \end{aligned} \tag{2.23}$$

However, the customary way to view the momentum operator is not in momentum space, but as a derivative in coordinate space. Expanding $|p\rangle$ in terms of coordinate space states, one can use completeness to express the $\langle\phi|\mathcal{P}|\psi\rangle$ in terms of wave functions,

$$\begin{aligned} \langle\phi|\mathcal{P}|\psi\rangle &= \int dxdy\frac{dp}{2\pi\hbar}\langle\phi|x\rangle\langle x|p\rangle p\langle p|y\rangle\langle y|\psi\rangle \\ &= \int dxdy\frac{dp}{2\pi\hbar}\langle\phi|x\rangle e^{ip(x-y)/\hbar}p\langle y|\psi\rangle \end{aligned} \tag{2.24}$$

The factor p can be changed into a derivative of the phase $e^{-ipy/\hbar}$ with respect to y,

$$\begin{aligned} \langle\phi|\mathcal{P}|\psi\rangle &= \int dxdy\frac{dp}{2\pi\hbar}\left(i\hbar\partial_y e^{ip(x-y)/\hbar}\right)\langle\phi|x\rangle\langle y|\psi\rangle \\ &= \int dxdy\frac{dp}{2\pi\hbar}\, e^{ip(x-y)/\hbar}\langle\phi|x\rangle\left(-i\hbar\partial_y\langle y|\psi\rangle\right). \end{aligned} \tag{2.25}$$

One can now perform the integral over p to get a factor of $\delta(x-y)$, which then allows one to integrate over y,

$$\begin{aligned} \langle\phi|\mathcal{P}|\psi\rangle &= \int dx\,\langle\phi|x\rangle\left(-i\hbar\partial_x\right)\langle x|\psi\rangle \\ &= \int dx\,\phi^*(x)\left(-i\hbar\partial_x\right)\psi(x), \\ \mathcal{P} &= \int dx\,|x\rangle\left(-i\hbar\partial_x\right)\langle x|. \end{aligned} \tag{2.26}$$

This algebra also shows that the one can express the momentum operator as a left-acting derivative, but with a change of sign.

Thus, the momentum operator can be thought of a $-i\hbar\partial/\partial x$. With some algebra, one can see that commuting $\mathcal{P}$ with $\mathcal{X}$ yields,

$$\begin{aligned} [\mathcal{P},\mathcal{X}] &= \int dx|x\rangle(-i\hbar\partial_x)(x\langle x|) - \int dx\,|x\rangle(x(-i\hbar\partial_x))\langle x| \\ &= -i\hbar\int dx\,|x\rangle\langle x| \\ &= -i\hbar\mathbb{I}. \end{aligned} \tag{2.27}$$

One should keep in mind that in many applications the definition of a momentum state in Eq. (2.12) is replaced with

$$\langle x|p\rangle \equiv \begin{cases} e^{ipx}/\sqrt{L}, & 0 < x < L \\ 0, & \text{otherwise} \end{cases}, \tag{2.28}$$

where $L \to \infty$. For this definition, $\langle x|p\rangle$ has dimensions of length$^{-1/2}$, and the values of p are discrete. The normalizations of the momentum states are then

$$\langle p|p'\rangle = \delta_{pp'}, \tag{2.29}$$

which are Kronecker delta functions instead of Dirac delta functions. These states are not actually eigenstates of the momentum operator because $-i\hbar\partial_x\langle x|p\rangle \neq p\langle x|p\rangle$ when $x < 0$ or $x > L$. Typically, this choice is useful when one is considering emission into a continuum of momentum states, and the density of such states factors into the answer. Because the density of states is proportional to L, the arbitrarily large lengths typically cancel from the answer. In practice, when reading texts or literature, or the later chapters of these notes, the choice is usually not explicitly stated, and the reader must discern the choice of convention based on whether factors of the length or volume appear.

Aside: Schrödinger's Wave Equation as an Operator

Before we find solutions, we regress to consider the notation and vocabulary associated with Schrödinger's equation. One often writes the Hamiltonian in this way,

$$H\psi(x) = E\psi(x), \quad H = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x). \tag{2.30}$$

Referring to the Hamiltonian H as an operator is actually incorrect, as it does not represent the fact that the Hamiltonian is an operator and should be sandwiched between a bra and ket. Operators are defined by matrix elements, and should referenced by two labels. In this case $\langle x'|H|x\rangle$, which should be considered a function of both x and x'. More correctly, beginning with H defined as an operator, then using completeness,

$$\begin{aligned} H|\psi\rangle &= E|\psi\rangle \\ \int dx'\langle x|H|x'\rangle\langle x'|\psi\rangle &= E\langle x|\psi\rangle. \end{aligned} \tag{2.31}$$

Now, one assumes that H does not mix $|x\rangle$ with $|x'\rangle$ unless x and x' are a the same position (or neighboring positions in the limit that one discretizes coordinate space). In that case one can define a function $h(x)$ as,

$$\langle x|H|x'\rangle = h(x)\delta(x - x'), \tag{2.32}$$

and Schrödinger's equation becomes

$$\begin{aligned} h(x)\langle x|\psi\rangle &= E\langle x|\psi\rangle, \\ h(x)\psi(x) &= E\psi(x), \\ h(x) &= -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x). \end{aligned} \tag{2.33}$$

This is not simply an exercise in notation. It emphasizes that the Hamiltonian operator does not mix states with x far from x'. Without such a constraint, one would not have local current conservation. Local conservation of current requires that charge cannot instantaneously transport across non-zero distances.

2.3 Local Charge Conservation and the Equation of Continuity

The equation of continuity is omnipresent throughout physics,

$$\partial_t \rho(\vec{r},t) = -\nabla\cdot\vec{j}(\vec{r},t). \tag{2.34}$$

To see that this infers that charge conservation is local, one applies the divergence theorem. For any volume V,

$$\begin{aligned}\partial_t \int_V d^3r\ \rho(\vec{r},t) &= -\int_V d^3r\ \nabla\cdot\vec{j}(\vec{r},t) \\ \frac{d}{dt}Q &= -\oint d\vec{A}\cdot\vec{j}(\vec{r},t),\end{aligned} \tag{2.35}$$

where Q is the charge within the volume V. By considering arbitrarily small volumes, this shows that the change in charge within that volume can be identified with the current density $\vec{j}$ exiting the surface of the volume. Thus, charge cannot leave a small volume in East Lansing, only to instantaneously appear in Kalamazoo. Instead, it appears in the neighboring volume, and might then make its way to Kalamazoo over time.

Schrödinger's equation allows one to identify both the charge and current densities. First, consider Schrödinger's wave equation and its complex conjugate,

$$\begin{aligned}\frac{-\hbar^2\nabla_2^2}{2m}\psi^*(\vec{r}_1,t_1)\psi(\vec{r}_2,t_2) + V(\vec{r}_2,t_2)\psi^*(\vec{r}_1,t_1)\psi(\vec{r}_2,t_2) &= -i\hbar\frac{\partial}{\partial t_2}\psi^*(\vec{r}_1,t_1)\psi(\vec{r}_2,t_2), \\ \frac{-\hbar^2\nabla_1^2}{2m}\psi^*(\vec{r}_1,t_1)\psi(\vec{r}_2,t_2) + V(\vec{r}_1,t_1)\psi^*(\vec{r}_1,t_2)\psi(\vec{r}_2,t_2) &= i\hbar\frac{\partial}{\partial t_1}\psi^*(\vec{r}_1,t_1)\psi(\vec{r}_2,t_2).\end{aligned} \tag{2.36}$$

The second equation is found by taking the complex conjugate of the first equation, then switching $(1\leftrightarrow 2)$. Subtracting the two equations, and using the fact that $\nabla_1^2-\nabla_2^2 = (\nabla_1-\nabla_2)\cdot(\nabla_1+\nabla_2)$,

$$\begin{aligned}&-\frac{\hbar^2}{2m}(\nabla_1+\nabla_2)\cdot(\nabla_1-\nabla_2)\left[\psi^*(\vec{r}_1,t_1)\psi(\vec{r}_2,t_2)\right] \\ &+\left[V(\vec{r}_1,t)-V(\vec{r}_2,t)\right]\left[\psi^*(\vec{r}_1,t_1)\psi(\vec{r}_2,t_2)\right] \\ &\qquad = -i\hbar\left(\frac{\partial}{\partial t_1}+\frac{\partial}{\partial t_2}\right)\left[\psi^*(\vec{r}_1,t_1)\psi(\vec{r}_2,t_2)\right].\end{aligned} \tag{2.37}$$

Then dividing by $i\hbar$,

$$\begin{aligned}&(\nabla_1+\nabla_2)\cdot\frac{-i\hbar}{2m}(\nabla_2-\nabla_1)\psi^*(\vec{r}_1,t_1)\psi(\vec{r}_2,t_2) \\ &+\frac{-i}{i\hbar}[V(\vec{r}_1,t)-V(\vec{r}_2,t)]\left[\psi^*(\vec{r}_1,t_1)\psi(\vec{r}_2,t_2)\right] \\ &\qquad = \left(\frac{\partial}{\partial t_1}+\frac{\partial}{\partial t_2}\right)\psi^*(\vec{r}_1,t_1)\psi(\vec{r}_2,t_2).\end{aligned} \tag{2.38}$$

Setting $\vec{r}_1 = \vec{r}_2$ and $t_1 = t_2$, one finds

$$\begin{aligned} \partial_t \rho(\vec{r}, t) &= -\nabla \cdot \vec{j}(\vec{r}, t), \\ \rho(\vec{r}, t) &= \psi^*(\vec{r}_1, t_1)\psi(\vec{r}_2, t_2) \\ \vec{j}(\vec{r}, t) &= -\frac{i\hbar}{2m}\psi^*(\vec{r}, t)\nabla\psi(\vec{r}, t) + \frac{i\hbar}{2m}[\nabla\psi^*(\vec{r}, t)]\psi(\vec{r}, t). \end{aligned} \tag{2.39}$$

If one associates $i\hbar\vec{\nabla}/m$ with the velocity, then the current looks like the velocity density. For Schrödinger's equation, the net number of particles is a conserved quantity. If the described particle has some charge q, then the charge density is $q\psi^*\psi$.

2.4 Potential Problems in One Dimension

Although few physics applications are in one dimension, many problems reduce to one dimension due to symmetry. For example, once there is rotational invariance, a three-dimensional solution to the Schrödinger wave function can be factored into products of a radial and angular pieces. The radial equation can then be mapped into a simple one-dimensional equation with an effective centrifugal potential. One-dimensional solutions form the foundation for understanding a wide range of physics beyond one-dimension.

Consider Schrödinger's wave equation in one dimension,

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi(x, t) + V(x)\Psi(x, t) = i\hbar\frac{\partial}{\partial t}\Psi(x, t). \tag{2.40}$$

If $\Psi(x, t)$ is an eigenstate of the Hamiltonian with eigenvalue E, the solution becomes

$$\Psi(x, t) = e^{-iEt/\hbar}\psi(x), \tag{2.41}$$

and the time derivative $i\hbar\partial/\partial t$ can be replaced by E in Schrödinger's wave equation. Because the time dependence is then simply a phase factor, all observables become fixed in time for eigenstates. For many applications, we will consider the steady-state case by looking for eigenstates. However, the fact that these are eigenstates does not necessarily imply that the currents are zero. An example is a momentum eigenstate. One can set boundary conditions such that current enters and exits the boundary, thus resulting in a steady-state solution, but one where the charge moves.

To find solutions for problems where the potential is continuous, it is sufficient to find solutions of the wave equation that have the correct behavior at $x \to \pm\infty$. If the potential is discontinuous at certain points, boundary conditions must be enforced at every point where a discontinuity is formed.

If the discontinuity is finite (not a delta function or an infinite potential wall), the boundary conditions at the discontinuity are:

1. $\psi(x)$ is continuous.
2. $\partial_x\, \psi(x)$ is continuous.

If ψ were discontinuous, the derivative would be infinite at that point and the current would be infinite at the discontinuity, which would make it impossible to enforce the equation of continuity. If the first derivative were discontinuous at some point y, the second derivative would be infinite at that point. To understand the second boundary conditions we integrate Schrödinger's equation from $x = y - \epsilon$ to $x = y + \epsilon$, where $\epsilon \to 0$. This yields

$$-\frac{\hbar^2}{2m}\left(\frac{\partial}{\partial x}\psi(x)|_{y+\epsilon} - \frac{\partial}{\partial x}\psi(x)|_{y-\epsilon}\right) = \int_{y-\epsilon}^{y+\epsilon} dx(E - V(x))\psi(x). \tag{2.42}$$

For finite potentials, the integrand in the integral of the r.h.s. of this equation is manifestly finite, thus as ϵ goes to zero, the r.h.s. must vanish, which demonstrates the boundary condition. One should note that for an infinite potential step the boundary conditions can be met by simply having the wave function vanish. In this case it would appear the slope changes suddenly, but because the wave function is zero there is no energy in the infinitesimal region.

Example 2.1: Sudden Disintegration of Square Well
Consider the simple potential,

$$V(x) = \begin{cases} \infty, & x < -a \\ 0, & -a < x < a \\ \infty, & x > a \end{cases}$$

The ground state wave function is

$$\psi(x) = \sqrt{\frac{1}{a}}\cos(\pi x/2a), \quad E = \frac{\hbar^2}{2m}\left(\frac{\pi}{2a}\right)^2.$$

If the well suddenly dissolves, what is the probability the particle will be observed with momentum p?
Solution: From Eq. (2.20), the probability of observing p is

$$\begin{aligned}
\frac{dP_\psi}{dp} &= \frac{1}{2\pi\hbar}|\langle p|\psi\rangle|^2, \\
\langle p|\psi\rangle &= \int dx\ \langle p|x\rangle\langle x|\psi\rangle \\
&= \int dx\ e^{-ipx/\hbar}\psi(x) \\
&= \frac{1}{\sqrt{a}}\int_{-a}^{a} dx\ [\cos(k) - i\sin(kx)]\cos(k_0 x)] \\
&= \frac{1}{\sqrt{a}}\int_{-a}^{a} dx\ \cos(kx)\cos(k_0 x), \\
k_0 \equiv \frac{\pi}{2a}, &\quad k \equiv \frac{p}{\hbar}.
\end{aligned}$$

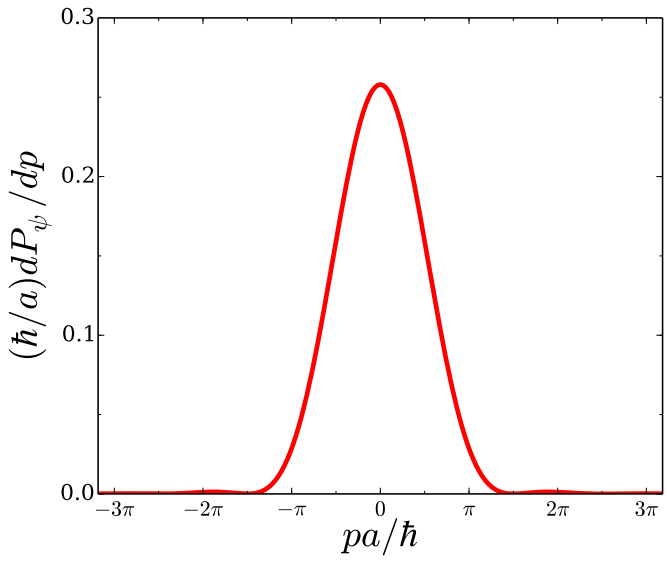


Figure 2.1: Differential probability of observing particles of momentum p from a suddenly disintegrating square well of width $2a$, solved in Example 2.1.

The term with $\sin(kx)$ was discarded by symmetry. Integrating,

$$\begin{aligned}\langle p|\psi\rangle &= \frac{1}{2\sqrt{a}}\int_{-a}^{a} dx\,\{\cos[(k+k_0)x]+\cos[(k-k_0)x]\}\\ &= \frac{1}{\sqrt{a}}\left\{\frac{\sin[(k+k_0)a]}{k+k_0}+\frac{\sin[(k-k_0)a]}{k-k_0}\right\}\\ &= \sqrt{a}\cos(ka)\left\{\frac{1}{ka+\pi/2}-\frac{1}{ka-\pi/2}\right\}\\ &= \sqrt{a}\cos(ka)\frac{\pi}{(ka)^2-\pi^2/4}.\end{aligned}$$

The dimension of $|\langle p|\psi\rangle|^2$ is that of length, and given that the units of $\hbar$ are momentum$\times$length (or energy$\times$time), the probability dP_ψ/dp indeed has units of inverse momentum. The probability is displayed in Fig. 2.1.

For many exercises a potential is chosen that is constant within some range of x, e.g. $0 < x < L$. In that region the form of the wave function is either sinusoidal or exponential, depending or whether the energy is above or below the potential If the energy is above the potential, i.e. the kinetic energy would appear positive, the form of the wave function is sinusoidal,

$$\psi(x) = Ae^{ikx} + Be^{-ikx}, \tag{2.43}$$

or equivalently

$$= C\cos(kx) + D\sin(kx).$$

The wave number k is chosen to match the kinetic energy,

$$\frac{\hbar^2k^2}{2m} = \mathrm{KE} = E - V(x). \tag{2.44}$$

If the energy is below the potentia., i.e. the kinetic energy would be negative, the form is exponential, or equivalently in terms of hyperbolic sines and cosines,

$$\psi(x) = Ae^{-\kappa x} + Be^{\kappa x}, \tag{2.45}$$
$$\textbf{or equivalently}$$
$$= C\cosh(\kappa x) + D\sinh(\kappa x).$$

The exponential scale κ is again chosen according to the energy

$$\frac{\hbar^2\kappa^2}{2m} = V(x) - E. \tag{2.46}$$

For the latter case the particles are energetically disallowed, but can penetrate into the disallowed region with an exponential form. In both cases, the constants A, B, C or C can be complex coefficients.

Example 2.2: Finite Potential Well
Consider the potential,

$$V(x) = \begin{cases} \infty, & x < 0 \\ -V_0, & 0 < x < a \\ 0, & x > a \end{cases}$$

Solve for the binding energy and wave function of the lowest energy state:
Solution: Assume the energy is negative. In the region of the well solutions of Schrödinger's equation are sines and cosines with wave number, $k = \sqrt{(2m(V_0 - B)/\hbar^2)}$, where B is the binding energy (a positive number between zero and V_0). The boundary condition at the origin is that the wave function must go to zero due to the infinite potential. Thus, only the sine piece remains in region I ($0 < x < a$),

$$\psi_I(x) = \sin(kx).$$

We have chosen an arbitrary normalization constant of unity.
In the second region, ($x > a$), the potential is zero and $E - V < 0$, so the solutions are exponentially growing or decaying. They are characterized by a decay constant of $q = \sqrt{(2mB/\hbar^2)}$. The exponentially growing piece can be thrown out as we wish to find a solution where some probability is near the origin.

$$\psi_{II}(x) = Ae^{-qx}$$

Thus far, we have neglected the two boundary conditions at $x = a$, which are necessary to determine the two unknowns, the binding energy B and the normalization factor A. Writing the two boundary conditions,

$$\sin(ka) = Ae^{-qa} \tag{2.47}$$
$$k\cos(ka) = -qAe^{-qa}$$

Dividing the two boundary conditions eliminates A and gives the relation,

$$\tan(ka) = \frac{-k}{q},$$
$$k = \sqrt{2m(V_0 - B)/\hbar^2},\ \ q = \sqrt{2mB/\hbar^2}.$$

This is a transcendental equation for B (or equivalently for k or q) which can be solved graphically, or on a computer, and keeping in mind that k must be between zero and $k_{\max} = \sqrt{2mV_0/\hbar^2}$. In order for the two functions above to intersect and have the same slope, the wave function must have a negative slope, and because $k < k_{\max} = \sqrt{2mV_0/\hbar^2}$, this can only happen when $k_{\max}a > \pi/2$. Thus, the solution disappears if the depth, V_0, is too small or if the the width a is too narrow. At the point where $k_{\max}a = \pi/2$, the solution has binding energy zero and $q \to 0$. Using the expression for $\tan(ka)$ one can see that $\tan(ka) = \infty$, or $ka = \pi/2$ as just stated.
In one dimension, if a potential remains zero or below zero for all x, there is always at least one bound state even if the potential depth or range become infinitesimally small. However, in this case the potential is positive, and in fact infinite, for $x < 0$, and the potential needs to have sufficient strength and extent to result in a bound state.

Example 2.3: Plane Wave Incident on Finite Barrier

Consider a plane wave of wave number k incident on a positive barrier,

$$V(x) = \begin{cases} 0, & x < 0 \\ V_0, & x > 0 \end{cases}.$$

Find the reflection and transmission probabilities for a particle with energy E which is greater than V_0.
Solution: In region I ($x < 0$) the solution consists of both an incoming, $\propto e^{ikx}$, and a reflected, $\propto e^{-ikx}$, component. Whereas in region II ($x > 0$) one only has a component moving in the positive x direction.

$$\psi_I(x) = e^{ikx} + Ae^{-ikx},\ k = \sqrt{2mE/\hbar^2},$$
$$\psi_{II}(x) = Be^{iqx},\ q = \sqrt{2m(E - V_0)/\hbar^2}.$$

The boundary conditions yield

$$1 + A = B,$$
$$k(1 - A) = qB.$$

Here, k and q are known, and the unknowns are A and B. Solving for A and B,

$$A = \frac{k - q}{k + q}, \quad B = \frac{2k}{k + q}$$

Thus, one quickly obtains the reflection and transmission probability.
In order to check the answer, one can see that the outgoing fluxes sum to the incoming flux.

$$\frac{\hbar k}{m}|A|^2 + \frac{\hbar q}{m}|B|^2 = \frac{\hbar k}{m}.$$

Fluxes are given by the square of the amplitude, which represents the probability density, multiplied by the velocity, or in this case the momentum divided by the mass.

We now consider a discontinuity at y, and perform the integral over an infinitesimal range centered about y. From Eq. (2.42) one could see that if the potential is finite in the neighborhood of y one obtains

$$-\frac{\hbar^2}{2m}\left(\frac{\partial}{\partial x}\psi(x)|_{y+\epsilon} - \frac{\partial}{\partial x}\psi(x)|_{y-\epsilon}\right) = 0. \tag{2.48}$$

However, if the potential is a delta function, integrating the right-hand side of Eq. (2.42) over the infinitesimal range can yield a finite value. Thus, if

$$V(x) = \beta\delta(x-y), \tag{2.49}$$

the boundary condition becomes

$$-\frac{\hbar^2}{2m}\left(\frac{\partial}{\partial x}\psi(x)|_{y+\epsilon} - \frac{\partial}{\partial x}\psi(x)|_{y-\epsilon}\right) = -\beta\psi(y). \tag{2.50}$$

The slope of the wave function is then not continuous at the delta function discontinuity.

Example 2.4: Delta Function Potential
Consider the potential,

$$V(x) = -\beta\delta(x), \;\; \beta > 0.$$

Find the binding energy of the bound state.
Solution: Assume the existence of a bound state of binding energy B. The solutions are exponentials with decay constant $k = \sqrt{2mB/\hbar^2}$, and the requirements that they go to zero at $x = \pm\infty$, and that they are equal at $x = 0$ demands

$$\psi(x) = \begin{cases} \exp(-kx) & x > 0 \\ \exp(kx) & x < 0 \end{cases}.$$

Plugging them into the boundary condition at $x = 0$, Eq. (2.50), gives

$$\frac{\hbar^2 k}{m} = \beta. \tag{2.51}$$

The binding energy is

$$B = \frac{\hbar^2 k^2}{2m} = \frac{\beta^2 m}{2\hbar^2},$$

and a bound state exists for any value of β.

2.5 Wave Packets

A plane wave confined to a finite region can be written as

$$\psi_{p,L}(x) = \begin{cases} 0, & x < 0 \\ \langle x|p,L\rangle = \frac{1}{\sqrt{L}}e^{ipx/\hbar}, & 0 < x < L \\ 0, & x > L \end{cases} \tag{2.52}$$

which is normalized in the length L, but is not really an eigenstate of the momentum operator due to the sharp cutoff at the boundaries. A more physical description of a wave that is confined to a finite region is a wave packet, which confines the probability to a finite region, and because of the uncertainty principle is accompanied by a spread in the momentum. This state $|Q\rangle$ will NOT be an eigenstate of the momentum operator, but will instead be a linear combination of states $|p\rangle$, which are eigenstates of the momentum operator $\mathcal{P} = -i\hbar\nabla$.

$$\begin{aligned} |Q\rangle &= \int \frac{dp}{2\pi\hbar} g(p-Q)|p\rangle, \\ g(p-Q) &= \langle p|Q\rangle. \end{aligned} \tag{2.53}$$

For our purposes we will assume g to be of a Gaussian form,

$$g(p-Q) = \alpha \exp\left(\frac{-(p-Q)^2}{4\Delta^2}\right). \tag{2.54}$$

As an exercise, one should check that this resulting wave packet is properly normalized by

$$\alpha = (2\pi)^{1/4}\sqrt{\frac{\hbar}{\Delta}}. \tag{2.55}$$

With this definition, the width of the wave packet in momentum space is Δ, i.e.

$$\begin{aligned} \langle Q|(\mathcal{P}-\bar{P})^2|Q\rangle &= \frac{1}{2\pi\hbar}\int dp\ (p-\bar{P})^2[g(p-Q)]^2 \\ &= \Delta^2, \\ \bar{P} &\equiv \langle Q|\mathcal{P}|Q\rangle. \end{aligned} \tag{2.56}$$

To understand the spatial shape of the packet, one can see that

$$\begin{aligned} \psi_Q(x) = \langle x|Q\rangle &= \int \frac{dp}{2\pi\hbar}\langle x|p\rangle\langle p|Q\rangle \\ &= \int \frac{dp}{2\pi\hbar}\, e^{ipx/\hbar} g(p-Q) \\ &= \frac{\alpha\Delta}{\hbar}\sqrt{\frac{2}{\pi}} e^{iQx/\hbar} e^{-x^2\Delta^2/\hbar^2}, \\ |\psi_Q(x)|^2 &= \frac{2\Delta}{\hbar}\sqrt{\frac{1}{2\pi}}\exp\left\{-\frac{(2\Delta)^2}{\hbar^2}\frac{x^2}{2}\right\}, \end{aligned} \tag{2.57}$$

which is also of Gaussian form, with the spatial spread being $\hbar/(2\Delta)$.

Thus, the product of the spread in momentum space multiplied by the spread in coordinate space is $\hbar/2$, exactly the minimum allowed by the uncertainty principle. If the packet had been described with $e^{ipx\hbar} \to e^{ip(x-x_0)/\hbar}$, the packet would be centered at x_0 instead of the origin.

For the wave packet at arbitrary times,

$$\psi_Q(x,t) = \int \frac{dp}{2\pi\hbar} e^{-iE_p t/\hbar + ipx/\hbar} g(p-Q), \tag{2.58}$$

we expect the packet to move in time. The packet has contributions from all momenta which represent differential additions with a wide variety of phases for each p. At most values of x these phases vary to the degree that the resulting wave function is near zero. However, at any given time t there is a point x for which the phases are constant as p is varied near Q. At this point, the differential contributions add in-phase and the wave function in coordinate space is a maximum. To find that point we take a derivative of the phase with respect to p at Q, and require it to be zero.

$$\frac{d}{dp}(-iE_p t + ipx)|_{p=Q} = 0, \tag{2.59}$$

$$x = \frac{dE_p}{dp}t.$$

Given that $dE/dp|_{p=Q} = v$, even for relativistic particles, one sees that

$$x = vt \tag{2.60}$$

which is not surprising. Here, $g(p)$ was assumed to be a real function, but if it were given a momentum-dependent phases, $e^{i\phi(p)}$, then the same condition would lead to

$$x = vt - \hbar d\phi/dp|_{p=Q}, \tag{2.61}$$

and the additional term would set the position of the wave packet away from the origin, even at $t = 0$.

Example 2.5: Wave Packet Reflecting off Finite Barrier
As an example where wave packets are involved, we will consider a packet as described above, incident on a potential barrier. The goal is to find the time delay of scattering. We consider a simple square well potential of the form,

$$V(x) = \begin{cases} 0, & x > 0 \\ V_0, & x < 0 \end{cases}$$

We also assume that the packet is narrow and that V_0 is larger than E_k, so the wave packet should completely reflect (assuming that $g(p-k)$ is narrow).
The incoming wave packet can have the form,

$$\psi_{\text{in}}(x,t) = \frac{1}{2\pi\hbar}\int dp\, e^{-iE_p t/\hbar - ipx/\hbar} g(p-k).$$

Such a packet should have a reflection, with the reflected wave packet,

$$\psi_{\text{out}}(x,t) = -\frac{1}{2\pi\hbar}\int dp\, e^{-iE_p t/\hbar + ipx/\hbar} e^{+2i\delta(p)} g(p-k).$$

The negative sign is chosen so that in the limit of an infinite potential $\delta(p) = 0$. The reflected packet has the same amplitude, because the flux must be the same both ways given that the reflection is total. However, there could be a phase factor that might be momentum dependent. The factor of two in the phase is a convention which we will encounter again when we discuss scattering theory.

Solving for the phase shift, we assume the solutions in regions I ($x > 0$) and II ($x < 0$) are

$$\psi_I(x) = e^{-ipx/\hbar} - e^{ipx/\hbar + 2i\delta(p)},$$
$$\psi_{II}(x) = Be^{qx/\hbar}, \;\; q = \sqrt{2m(V_0 - E)}.$$

The boundary conditions give

$$1 - e^{2i\delta} = B$$
$$-ip(1 + e^{2i\delta}) = qB.$$

The solution for the phase shift is

$$\tan\delta = \frac{p}{q}.$$

Looking for the point x where the phase is stationary, one finds the expression

$$x = vt - 2\frac{d}{dp}\delta(p).$$

Using $p^2 + q^2 = 2mV_0$, one finds

$$\frac{d\delta}{dp} = \frac{2mV_0}{q^3}.$$

From the equation for ψ_{out}, one can see that the condition for a stationary phase in the outgoing wave packet becomes

$$\frac{d}{dp}\left\{-iE_pt/\hbar + ipx/\hbar + 2i\delta(p)\right\} = 0, \tag{2.62}$$

$$\Delta x = -2\hbar\frac{d}{dp}\delta(p),$$
$$\Delta t = 2\frac{\hbar}{v_p}\frac{d}{dp}\delta(E) = 2\hbar\frac{d}{dE}\delta(E).$$

This time delay is relative to the case where $V_0 = \infty$. The same ideas will be applied to spherical waves when we cover scattering theory, where the phase shifts, δ, provide a standard, and insightful, means for understanding the effects of a potential in scattering.

2.6 The Harmonic Oscillator

The harmonic oscillator provides the basis for much of physics. In fact, quantum field theory considers every point in space to have it's own oscillators for every type of quantum field. We begin with a single oscillator with the Hamiltonian,

$$H = \frac{\mathcal{P}^2}{2m} + \frac{m\omega^2}{2}\mathcal{X}^2, \tag{2.63}$$

where the spring constant is expressed in terms of a frequency ω, $k = m\omega^2$. Here, $\mathcal{X}$ and $\mathcal{P}$ are operators.

Dirac's solution to the problem involves first defining two new operators,

$$a \equiv \sqrt{\frac{m\omega}{2\hbar}}\mathcal{X} + i\sqrt{\frac{1}{2\hbar m\omega}}\mathcal{P}, \quad a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\mathcal{X} - i\sqrt{\frac{1}{2\hbar m\omega}}\mathcal{P}. \tag{2.64}$$

The operators are known as the annihilation (or destruction) and creation operators respectively for reasons to be seen below. In some texts, these operators are referred to as ladder operators.

The operators satisfy simple commutation relations,

$$[a, a^\dagger] = 2i\sqrt{\frac{m\omega}{2\hbar}}\sqrt{\frac{1}{2\hbar m\omega}}[\mathcal{P}, \mathcal{X}] = 1. \tag{2.65}$$

Furthermore, by substituting the expressions for $\mathcal{X}$ and $\mathcal{P}$ in terms of a and $a^\dagger$,

$$\begin{aligned} \mathcal{X} &= \sqrt{\frac{\hbar}{2m\omega}}(a + a^\dagger), \\ \mathcal{P} &= -i\sqrt{\frac{\hbar m\omega}{2}}(a - a^\dagger), \end{aligned} \tag{2.66}$$

into the Hamiltonian, Eq. (2.63), the Hamiltonian may be expressed as

$$\begin{aligned} H &= \frac{1}{2m}\left[-i\sqrt{\frac{\hbar m\omega}{2}}(a - a^\dagger)\right]^2 + \frac{m\omega^2}{2}\left[\sqrt{\frac{\hbar}{2m\omega}}(a + a^\dagger)\right]^2 \\ &= \frac{\hbar\omega}{4}\left[-a^2 + aa^\dagger + a^\dagger a - (a^\dagger)^2 + a^2 + aa^\dagger + a^\dagger a + (a^\dagger)^2\right] \\ &= \frac{\hbar\omega}{2}(aa^\dagger + a^\dagger a) \\ &= \hbar\omega\left(a^\dagger a + \frac{1}{2}\right). \end{aligned} \tag{2.67}$$

To see that the creation operator does exactly what it sounds like it should do, consider an eigenstate of the Hamiltonian such that

$$a^\dagger a|n\rangle = n|n\rangle. \tag{2.68}$$

At this point we do not know that the eigenvalues, n, of $a^\dagger a$ are integers. By using the commutation relations, $aa^\dagger = a^\dagger a + 1$, one can see that

$$\begin{aligned} a^\dagger a(a^\dagger|n\rangle) &= a^\dagger(a^\dagger a + 1)|n\rangle \\ &= (n+1)a^\dagger|n\rangle \end{aligned} \tag{2.69}$$

by commuting the $a^\dagger$ to the far left. Doing the same with the state $a|n\rangle$ one finds

$$\begin{aligned} a^\dagger a(a|n\rangle) &= (aa^\dagger - 1)a|n\rangle \\ &= (n-1)(a|n\rangle). \end{aligned} \tag{2.70}$$

Thus, to within a normalization constant,

$$a^\dagger|n\rangle \propto |n+1\rangle,\ \ a|n\rangle \propto |n-1\rangle. \tag{2.71}$$

To calculate the normalization constant Z,

$$\begin{aligned} a^\dagger|n\rangle &= \frac{1}{\sqrt{Z}}|n+1\rangle, \\ \langle n|aa^\dagger|n\rangle &= \frac{1}{Z} \\ &= \langle n|(a^\dagger a+1)|n\rangle = n+1, \\ Z &= \sqrt{\frac{1}{n+1}}, \\ a^\dagger|n\rangle &= \sqrt{n+1}|n+1\rangle. \end{aligned} \tag{2.72}$$

Similarly for the destruction operator,

$$\begin{aligned} a|n\rangle &= \frac{1}{\sqrt{Z}}|n-1\rangle, \\ \langle n|a^\dagger a|n\rangle &= n, \\ Z &= \frac{1}{\sqrt{n}}, \\ a|n\rangle &= \sqrt{n}|n-1\rangle. \end{aligned} \tag{2.73}$$

The operator $a^\dagger a$ is referred to as the number operator, and because the energy is expressed in terms of the number operator, and because that energy must not have arbitrarily small negative values, the values of n cannot be arbitrarily negative. Thus, the sequence of repeatedly applying the operator a to some state n must terminate at some point. This only happens if there is a state with $n = 0$,

$$a|n=0\rangle = 0, \tag{2.74}$$

which constrains all the n to be integers. With the constraint that n is an integer, the normalizations can be expressed as

$$(a^\dagger)^n|0\rangle = \sqrt{n!}|n\rangle. \tag{2.75}$$

and the eigenenergies are

$$\begin{aligned} E_n &= (n+1/2)\hbar\omega, \\ n &= 0,1,2,\cdots \end{aligned} \tag{2.76}$$

We will see later on that similar tricks are used with angular momentum raising and lowering operators.

Finally, it is straight-forward to find the ground state wave function if one is sufficiently prescient to guess that the form of the solution is a Gaussian with some as-yet-undetermined width b.

$$\psi_0(x) = e^{-x^2/(2b^2)}. \tag{2.77}$$

For the moment, we neglect the normalization. To show that this is a solution we first take derivatives with respect to x.

$$\frac{d}{dx}\psi_0(x) = -\frac{x}{b^2}\psi_0(x), \quad \frac{d^2}{dx^2}\psi_0(x) = \frac{x^2}{b^4}\psi_0(x) - \frac{1}{b^2}\psi_0(x). \tag{2.78}$$

Plugging this into the Schrödinger equation,

$$-\frac{\hbar^2}{2m}\left[\frac{x^2}{b^4} - \frac{1}{b^2}\right]\psi_0(x) + \frac{m\omega^2 x^2}{2}\psi_0(x) = E\psi_0(x). \tag{2.79}$$

allows one to determine b and E by inspection.

$$b = \sqrt{\frac{\hbar}{m\omega}}, \quad E = \frac{1}{2}\hbar\omega. \tag{2.80}$$

Indeed, the energy is what was expected from the earlier arguments with creation and destruction operators. One may also calculate the normalization Z by enforcing the constraint

$$\psi_0(x) = Z^{-1/2}e^{-x^2/(2b^2)}, \quad Z = \int_{-\infty}^{\infty} dx e^{-x^2/b^2}. \tag{2.81}$$

This gives $Z = (\pi^{1/2}b)$.

If one were to consider an n dimensional problem,

$$H\Psi = -\frac{\hbar^2}{2m}\frac{d^2}{dx_1^2}\Psi - \frac{\hbar^2}{2m}\frac{d^2}{dx_2^2}\Psi \cdots - \frac{\hbar^2}{2m}\frac{d^2}{dx_n^2}\Psi + \frac{m\omega^2}{2}x_1^2\Psi + \frac{m\omega^2}{2}x_2^2\Psi \cdots + \frac{m\omega^2}{2}x_n^2\Psi, \tag{2.82}$$

one would write the solution as

$$\Psi(x_1, x_2, \cdots x_n) = \psi_1(x_1)\psi_2(x_2)\cdots\psi_n(x_n), \tag{2.83}$$

where ψ_i are the solutions to the 1-d Schrödinger equation with eigenenergies $\hbar\omega(n_i + 1/2)$, and the total energy is

$$E = \hbar\omega\sum_i (n_i + 1/2), \tag{2.84}$$

with the ground state energy being for $n = 0$, $E_0 = \hbar\omega/2$. For a three-dimensional oscillator, one can consider the x, y and z motion as three separate oscillators, and the ground state energy would be $3\hbar\omega/2$.

Example 2.6: Practice with Creation and Destruction Operators
Evaluate the following matrix element:

$$\langle m|aa^\dagger a^\dagger aa|n\rangle.$$

Solution:

Using the facts that $a|n\rangle = \sqrt{n}|n-1\rangle$ and $a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$,

$$\begin{aligned}\langle m|aa^\dagger a^\dagger aa|n\rangle &= \sqrt{n}\langle m|aa^\dagger a^\dagger a|n-1\rangle \\ &= (n-1)\sqrt{n}\langle m|aa^\dagger|n-1\rangle \\ &= (n-1)n\langle m|a|n\rangle \\ &= (n-1)n^{3/2}\langle m|n-1\rangle \\ &= (n-1)n^{3/2}\delta_{m,n-1}.\end{aligned}$$

2.7 Propagators, Green's Functions and Integral Equations

A propagator is merely an incarnation of an evolution operator and is often confused with being the same as a Green's function, which it sometimes is. A propagator in coordinate space is defined as

$$K(x,t;x',t') \equiv \langle x|\exp\left[\frac{-iH(t-t')}{\hbar}\right]|x'\rangle, \tag{2.85}$$

and represents the amplitude from evolving from a position $|x'\rangle$ at time t' to a position $|x\rangle$ at time t. Rather than propagating from x' to x, one can also define propagators from some momentum state p' to some final momentum state p. Propagators form the basis of diagrammatic perturbation theory, and in that application provide a physical understanding of specific processes.

If the eigenstates, $|a\rangle$, of H are known the propagator may be written as

$$K(x,t;x',t') = \sum_a e^{-i\epsilon_a t/\hbar}\langle x|a\rangle\langle a|x'\rangle. \tag{2.86}$$

Because K is the evolution operator, one sees that determining the propagator is equivalent to solving the Schrödinger equation,

$$\psi(x,t) = \int d^3x' K(x,t;x',t')\psi(x',t'). \tag{2.87}$$

In fact, the propagator is a solution of Schrödinger's equation for $t > t'$ with the boundary condition that at t' the wave function is at x',

$$\left(-\frac{\hbar^2\nabla^2}{2m} + V(x)\right)K(x,t;x',t') = i\hbar\frac{\partial}{\partial t}K(x,t;x',t'), \tag{2.88}$$

while being zero for $t < t'$ and equal to $\delta^3(x-x')$ when $t = t_{0+}$.

We now consider the simple case of a free particle in one dimension. In that case the eigenstates of the Hamiltonian are momentum states and

$$\begin{aligned}K_0(x,t;x',t') &= \left(\frac{1}{2\pi\hbar}\right)\int dp \exp\left[\frac{ip\cdot(x-x')}{\hbar} - \frac{ip^2(t-t')}{2m\hbar}\right] \\ &= \sqrt{\frac{m}{2\pi i\hbar(t-t')}}\exp\left[\frac{im(x-x')^2}{2\hbar(t-t')}\right].\end{aligned} \tag{2.89}$$

The integral in the last step was performed by completing the square. Note that the phase in the exponential looks $(1/2)mv^2t/\hbar$, where the velocity is given by $\Delta x/\Delta t$. A similar expression can easily be found in higher dimension. To see that this form approaches $\delta(x-x')$ as $t \to t'$, one can check:

1. that $\int dx K(x,t;x',t') = 1$ for any time $t > t'$.
2. that at small time differences the phase oscillates infinitely quickly with varying x except when $x = x'$. This rapid variation of phases would cause any integral of the propagator over a finite interval to vanish unless that interval includes $x = x'$.

Adding a potential makes finding the propagator much more difficult. One can see that if the propagator is expressed as

$$K(x,t;x',t') = K_0(x,t;x',t') + \frac{1}{i\hbar}\int dt''dx'' K_0(x,t;x'',t'')V(x'',t'')K(x'',t'';x',t'), \tag{2.90}$$

the Schrödinger equation will be satisfied. This expression for the propagator can be viewed as the integral representation of Schrödinger's wave equation. Note that the second propagator in the right side is the "full" propagator which means that solving the equation is not simply a matter of performing an integral. Instead, one must find a self-consistent solution.

By replacing K with K_0 in the right-hand side of Eq. (2.90), one has found the "first-order perturbative corrections to the propagator". We will see this again later.

2.8 Phase Space Density and Wigner Functions

We have seen that $|\psi(x)|^2 = |\langle x|\psi\rangle|^2$ can be identified as the density dN/dx, and $|\psi(p)|^2 = |\langle p|\psi\rangle|^2$ can be identified as the number per $dp/(2\pi\hbar)$. One can also get a measure of the number per $dpdx$, which simultaneously expresses a probability of observing the particle with both a specific momentum and density by performing what is called a *Wigner transform* of the wave function,

$$f(p,x) = \int d\delta x\, \psi^*(x+\delta x/2)\psi(x-\delta x/2)e^{ip\delta x/\hbar}. \tag{2.91}$$

With this definition,

$$\begin{aligned}\int \frac{dp}{2\pi\hbar} f(p,x) &= \int d\delta x\, \psi^*(x+\delta x/2)\psi(x-\delta x/2)\int \frac{dp}{2\pi\hbar}\, e^{ip\delta x/\hbar} \\ &= \psi^*(x)\psi(x).\end{aligned} \tag{2.92}$$

Further, one can express $f(p,x)$ in terms of $\psi(p)$,

$$f(p,x) = \int d\delta x \int \frac{dp_1}{2\pi\hbar} e^{-ip_1(x+\delta x/2)/\hbar}\psi^*(p_1)\frac{dp_2}{2\pi\hbar}e^{ip_2(x-\delta x/2)/\hbar}\psi(p_2)e^{ip\delta x/\hbar}. \tag{2.93}$$

Integrating over δx,

$$f(p,x) = \int \frac{d\delta p}{2\pi\hbar}\psi^*(p+\delta p/2)\psi(p-\delta p/2)e^{-i\delta p x/\hbar}, \tag{2.94}$$

and integrating over x,

$$\int dx\, f(p,x) = \psi^*(p)\psi(p). \tag{2.95}$$

Thus, when one integrates over either x, as seen in Eq. (2.92), or if one integrates over p, as in Eq. (2.95), it appears that $f(p,x)$ behaves like the phase space density,

$$\frac{dN}{dpdx} = \frac{f(p,x)}{2\pi\hbar}. \tag{2.96}$$

This is called the phase space density, or occupancy, because it can be related to the number of particles per single-particle level. To demonstrate this, we consider the limit of a large box of length with an infinite square well potential confining the particle to within $0 < x < L$, the boundary conditions give solutions of

$$\begin{aligned} \psi_n(x) &= \sin(p_n x/\hbar) = \sin(n\pi x/L), \;\; n = 1,2,3,\cdots, \\ p_n &= n\pi\hbar/L. \end{aligned} \tag{2.97}$$

The number of states with momentum p within some range Δp is

$$\Delta N = \frac{L}{\pi\hbar}\Delta p. \tag{2.98}$$

If one uses plane waves, which allows the states p to be both positive or negative, one alters the expression to

$$\Delta N = \frac{L}{2\pi\hbar}\Delta p. \tag{2.99}$$

Because $f(p,x)$ gives the number of particles per unit $\Delta x\Delta p/(2\pi\hbar)$, and the number of states is $\Delta x\Delta p/(2\pi\hbar)$, $f(p,x)$ can also be viewed as the number of particles per single particle level. Of course, this interpretation is based on the idea that $\Delta x = L$ is large enough so that Δn can be treated as a continuous function.

The interpretation of $f(p,x)$ might seem to violate the uncertainty principle, which forbids the simultaneous measurement of p and x. However, simply having an expression for $f(p,x)$ does not mean that you can measure both p and x simultaneously. In fact $f(p,x)$ can be negative, or greater than unity, for some values of p and x, and for some wave functions. This disqualifies it as a probabilistic measure because probabilities must lie between zero and unity. It can only be identified probabilistically in the classical limit, where the averaging is over a sufficiently large range of p and x to include many quantum state. Nonetheless, even away from the classical limit, $f(p,x)$ has the property that if it is integrated over p one obtains the probabilistic density in coordinate space, and if one integrates over x, one finds the probabilistic density in momentum space.

Example 2.7: Wigner Transform of Harmonic Oscillator Wave Functions
Here, we find the phase space density, a.k.a. the Wigner transform, of the ground state of the

harmonic oscillator.

$$\begin{aligned}\psi_0(x) &= \frac{1}{\pi^{1/4}b^{1/2}}e^{-x^2/2b^2},\\ f(p,x) &= \frac{1}{\pi^{1/2}b}\int d\delta x\; e^{ip\delta x/\hbar}e^{-(x+\delta x/2)^2/2b^2}e^{-(x-\delta x/2)^2/2b^2}\\ &= \frac{1}{\pi^{1/2}b}e^{-x^2/b^2}\int d\delta x\; e^{ip\delta x/\hbar}e^{-\delta x^2/4b^2}\\ &= 2e^{-x^2/b^2-p^2b^2/\hbar^2}.\end{aligned}$$

One can check the normalization by showing that

$$\int \frac{dpdx}{2\pi\hbar}\, f(p,x) = 1.$$

2.9 Numerically Solving Schrödinger's Equation

Here, we provide algorithms for solving one-dimensional Schrödinger's equations for two types of problems. For both types we consider potentials that are infinite for $x < 0$. This boundary condition also applies for spherical wave decompositions in three dimensions, where for a particular spherical wave the problem reduces to a one-dimensional problem with an effective potential for angular momentum, but because there are no negative radii, one sets a boundary condition that $\psi(r = 0) = 0$. The first type of problems to be considered is solving for a bound state energy for a finite (for $x > 0$) potential that vanishes beyond some position, $V(x > a) = 0$, and the second type of problem involves finding the phase shift for the same kind of potential.

To find a bound state energy numerically, one must first guess at a binding energy B. For the sake of convenience, instead of B we will refer to q, where $\hbar^2q^2/2m = B$. For $x > a$, the bound state wave function behaves as Ae^{-qx}, where A is an unknown constant. For $x < a$, we solve the problem numerically on a grid of resolution Δx,

$$\psi(x = 0) = 0,\;\; \psi(x = \Delta x) = \psi_1. \tag{2.100}$$

Here, ψ_1 can be any value. Normalization will be chosen later. For the sake of convenience, one might choose a real value such as dx. Next, one uses the discrete representation of Schrödinger's equation,

$$\frac{-1}{(\Delta x)^2}\left(\psi_{n+1} - 2\psi_n + \psi_{n-1}\right) = \left(-q^2 - \frac{2mV(x_n)}{\hbar^2}\right)\psi_n, \tag{2.101}$$

to solve for ψ_{n+1},

$$\psi^{n+1} = 2\psi_n - \psi_{n-1} - (\Delta x)^2\left(-q^2 - \frac{2mV(x_n)}{\hbar^2}\right)\psi_n. \tag{2.102}$$

One can iterate this forward, first solving for ψ_2, then ultimately for ψ_N, where $N = a/dx$. At that point one can consider the boundary conditions. By comparing logarithmic derivatives, $(d\psi/\Delta x)/\psi$, one can ignore the constant A, and see that

$$\frac{\psi_{N+1} - \psi_{N-1}}{2(\Delta x)\psi_N} = -q. \tag{2.103}$$

The difficulty is finding the value of q that satisfies this boundary condition at $x = a$. This involves guessing a value, then comparing the equivalence. One then adjusts q until one finds a satisfactory fit. This can be done with Newton's method, or with a root-finding package. The bound state energy is then $E = -B = -\hbar^2 q^2/(2m)$.

The second class of problem to be considered is that of finding the phase shift for a given incoming wave number k. This is easier because, the energy is given. Again, discretize the wave function, and assume that the wave function behaves like the incoming wave, e^{-ikx},

$$\phi_{N+1} = e^{-ik(N+1)\Delta x}, \;\; \phi_N = e^{-ikN\Delta x}. \tag{2.104}$$

One then iterates forward,

$$\phi_{n-1} = 2\phi_n - \phi_{n+1} - (\Delta x)^2 \left(q^2 - \frac{2mV(x_n)}{\hbar^2} \right) \phi_n. \tag{2.105}$$

Ultimately, one finds ϕ_0, which of course does not satisfy the boundary condition. Because Schrödinger's equation is real, one also knows that ϕ^* is also a solution, also not satisfying the boundary condition. However, one can take the combination,

$$\psi_n = \phi_n - e^{2i\delta}\phi_n^*, \tag{2.106}$$

choosing the phase shift to fit the boundary condition,

$$e^{2i\delta} = \frac{\phi_0}{\phi_0^*}. \tag{2.107}$$

With this choice of δ, the net wave function at $x = 0$ vanishes. This method even works for cases where the potential blows up at the origin, although one may have to extrapolate the phase to $x = 0$ from $x = \Delta x$ and $x = 2\Delta x$.

2.10 Combining Coordinate Space and Spin

The overlap $\psi(x) = \langle x|\Psi\rangle$ describes the overlap of the a state $|\Psi\rangle$ with a state located at x. The wave function $\psi(x)$ can be thought of as having values at each point x. For spin-half, one represents the overlap with the $|\uparrow\rangle$ and $|\downarrow\rangle$ states with a two-component vector. Of course, states can vary in both position and spin, and the spin might be correlated with position. In that case, one must consider the overlap, $\langle s, x|\Psi\rangle$, where s refers to the spin. In general, it can be represented by a two-dimensional vector, with each component depending on x,

$$\langle s, x|\Psi\rangle = \begin{pmatrix} \psi_\uparrow(x) \\ \psi_\downarrow(x) \end{pmatrix}. \tag{2.108}$$

The functions $\psi_\uparrow$ and $\psi_\downarrow$ can be any functions as long as the overall wave function is normalized,

$$\langle\Psi|\Psi\rangle = \int dx \left(\psi^*_\uparrow\ (x)\psi_\uparrow\ (x) + \psi^*_\downarrow(x)\psi_\downarrow(x)\right) = 1. \tag{2.109}$$

If the position and spin information are independent, i.e. the upper and lower components have the same spatial wave function, one might represent the overall wave functions as

$$\langle s, x|\Psi\rangle = \psi(x)\chi, \tag{2.110}$$

where χ is some two component wave function. But, this last form is not as general at that in Eq. (2.108), which accommodates the possibility that the spin features varies with position. For particles of higher spin, one needs to increase the dimensionality of the vector. For example, for massive spin-one particles, there are three spin states, $m = -1, 0, 1$ and one needs to use three-dimensional vectors.

2.11 Exercises

1. Proof that $\hbar = 0$: Consider a normalized momentum eigenstate of the momentum operator $|q\rangle$, i.e. $\mathcal{P}|q\rangle = q|q\rangle$ and $\langle q|\mathcal{P} = \langle q|q$. Consider the expectation,

$$\begin{aligned}\langle q|(\mathcal{P}\mathcal{X} - \mathcal{X}\mathcal{P})|q\rangle &= \langle q|(q\mathcal{X} - \mathcal{X}q)|q\rangle \\ &= q\langle q|(\mathcal{X} - \mathcal{X})|q\rangle = 0.\end{aligned}$$

   However the commutation relation, $\mathcal{P}\mathcal{X} - \mathcal{X}\mathcal{P} = -i\hbar$, so we also have

$$\langle q|(\mathcal{P}\mathcal{X} - \mathcal{X}\mathcal{P})|q\rangle = -i\hbar.$$

   Comparing the two equations, $\hbar = 0$.

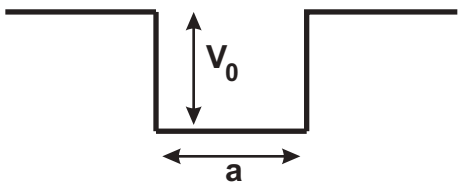
   What went wrong?

2. Prove that the average kinetic energy is always positive, i.e.

$$\langle -\frac{\hbar^2\partial_x^2}{2m}\rangle = -\frac{\hbar^2}{2m}\int dx\ \psi^*(x)\partial_x^2\psi(x) > 0.$$

3. Consider the one-dimensional potential,

$$V(x) = \begin{cases} 0, & x < -a \\ -V_0, & -a < x < a \\ 0, & x > a \end{cases}$$



   For fixed a, find the minimum V_0 for the number of bound states to equal or exceed 1,2,3....

4. Consider a particle of mass m under the influence of the potential,

$$V(x) = V_0\theta(-x) - \frac{\hbar^2}{2m}\beta\delta(x-a), \quad V_0 \to \infty,\ \beta > 0.$$

   (a) Find the transcendental equation for the energy of a bound state?
   (b) What is the minimum value of β for a ground state?
   (c) For increasing β can one find more than one bound state?

5. Consider a plane wave moving in the $-\hat{x}$ direction to be reflected off the delta function potential, For($x > a$) the plane wave will have the form

$$e^{-ikx} - e^{2i\delta}e^{ikx}.$$

   (a) Find the phase shift δ as a function of ka, and plot for $\beta a = 0.5$ and for $0 < ka < 10$. Because addition of $n\pi$ to the phase shift is arbitrary, translate all phases to angles between zero and π.
   (b) Repeat for $\beta a = 0.99, 1.01, 1.5$.

6. Consider a particle of mass m interacting with a repulsive δ function potential,

$$V(x) = \frac{\hbar^2}{2m}\beta\delta(x).$$

   Consider particles of energy E incident on the potential.

   (a) What fraction of particles are reflected by the potential?
   (b) Show that the currents for$x >$ and for $x < 0$ are the same.

7. Consider a three-dimensional harmonic oscillator with quantum numbers n_x, n_y and n_z. How many states are there with a given $N = n_x + n_y + n_z$? Find a closed expression (no sum). Test it for all $n \leq 3$.
8. Calculate $\langle 0|aaa^\dagger aa^\dagger a^\dagger|0\rangle$ and $\langle n|a^\dagger a^\dagger a^\dagger a|m\rangle$.
9. Find $\psi_1(x)$, the wave function of the first excited state by applying $a^\dagger$, defined in Eq. (2.64), to the ground state.
10. Consider a particle of mass m in a harmonic oscillator with spring constant $k = m\omega^2$.

    (a) Write the momentum and position operators for a particle of mass m in a harmonic oscillator characterized by frequency ω in terms of the creation and destruction operators.
    (b) Calculate $\langle n|\mathcal{X}^2|n\rangle$ and $\langle n|\mathcal{P}^2|n\rangle$. Compare the product of these two matrix elements to the constraint of the uncertainty relation as a function of n.
    (c) Show that the expectation value of the potential energy in an energy eigenstate of the harmonic oscillator equals the expectation value of the kinetic energy in that state.

11. (a) What is the representation of the position operator in the momentum basis – how is $\langle p|\mathcal{X}|\Psi\rangle$ related to $\langle p|\Psi\rangle$? Use the completeness relation, $\int dx|x\rangle\langle x| = \mathbb{I}$ and the fact that $\langle p|x\rangle = e^{-ipx/\hbar}$.

(b) Suppose that the potential is $v(\mathrm{x}) = (k/2)x^2$. What is the Schrödinger equation written in momentum space; i.e. what is the equation of motion of the amplitude $\langle p|\Psi(t)\rangle$?

12. Consider a potential

$$V(x) = \begin{array}{ll} 0, & x < -a \\ u(x), & -a < x < a \\ 0, & x > a \end{array}$$

where $u(x)$ is an arbitrary real function. Consider a wave incident from the left. Suppose that the transmission amplitude, defined as the ratio of the transmitted wave at $x = a$ to the incident wave at $x = -a$, is $S(E)$. Now consider a wave incident from the right. Show that the transmission amplitudes, $|S(E)|$, are the same for both directions. *(Hint: the Schrödinger equation in this case is a real equation, so the complex conjugate of a solution is also a solution.)*

13. Consider the creation and annihilation operators for the harmonic oscillator in the Heisenberg representation, $a^\dagger(t)$ and $a(t)$.

(a) Derive and solve the equations of motion for $a(t)$ and $a^\dagger(t)$ for the harmonic oscillator. I.e., first find $(d/dt)a(t)$ and $(d/dt)a^\dagger(t)$. Then write the solution of the differential equation given $a(t0 = 0) = a$ and $a^\dagger(t = 0) = a^\dagger$.

(b) Calculate $[a(t), a^\dagger(t')]$.

14. Calculate the correlation function $\langle 0|x(t)x(t')|0\rangle$ for the harmonic oscillator where $|0\rangle$ is the harmonic oscillator ground state, and $x(t)$ is the position operator in the Heisenberg representation. Hint: use the expressions for $a(t)$ and $a^\dagger(t)$ from the previous problem. Then solve for the equations of motion for both $x(t)$ and $p(t)$.

15. What are the matrix elements of the operator $1/|\vec{p}|$ in the position representation? That is, find

$$\langle \mathrm{r}|\frac{1}{|\mathrm{p}|}|\mathrm{r}'\rangle.$$

Work the problem in three dimensions.

16. Calculate the Wigner transform $f(p, x)$ for a particle in the ground state of an infinite square well potential,

$$V(x) = \begin{cases} \infty, & x < 0 \\ 0, & -a/2 < x < a/2 \\ \infty, & x > a \end{cases} .$$

Are there any regions with phase space densities either greater than unity or less than zero?

3 Charged Particles in Electromagnetic Fields

3.1 Vector and "Scalar" Potentials

Electric and magnetic fields are determined by the vector potential $\vec{A}$ and the scalar potential Φ,

$$\vec{E} = -\nabla\Phi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t}, \quad \vec{B} = \nabla \times \vec{A}. \tag{3.1}$$

The units of $e\Phi$ and $e\vec{A}$ are both energy. This is consistent with cgs units, where the magnetic and electric fields have the same units. The coupling e must change sign if the charge of the particle is reversed, i.e. this sign would be reversed for an electron.

To add electromagnetic interactions to the Schrödinger equation, one begins with a Hamiltonian where such interactions are missing,

$$i\hbar\partial_t = \frac{(-i\hbar\nabla)^2}{2m} + V(\vec{r}), \tag{3.2}$$

where $V(r)$ refers to potentials other than those from electromagnetism. Next, one performs "minimal substitution", which is to make the replacements

$$\begin{aligned} -i\hbar\nabla &\to -i\hbar\nabla - e\vec{A}/c, \\ i\hbar\partial_t &\to i\hbar\partial_t - e\Phi. \end{aligned} \tag{3.3}$$

Relativistically, the "scalar" potential is the zero$^{\text{th}}$ component of the four-vector, $A_0 = \Phi$, so the principal of minimal substitution becomes $i\hbar\partial_\mu \to i\hbar\partial_\mu - eA_\mu$. Thus,there is nothing "scalar" about the scalar potential, as it is not invariant under Lorentz transformations. Including the electromagnetic interaction, Schrödinger's wave equation becomes:

$$\frac{\left[-i\hbar\nabla - e\vec{A}(\vec{r},t)/c\right]^2}{2m}\psi(\vec{r},t) = \left[i\hbar\partial/\partial t - e\Phi(\vec{r},t)\right]\psi(\vec{r},t), \tag{3.4}$$

$$\begin{aligned} H\psi(\vec{r},t) &= i\hbar\partial_t\psi(\vec{r},t), \\ H &= \frac{1}{2m}\left(\vec{\mathcal{P}} - \frac{e\vec{A}}{c}\right)^2 + e\Phi(\vec{r},t) \\ &= \frac{1}{2m}\left[\mathcal{P}^2 - \frac{e}{c}(\vec{\mathcal{P}}\cdot\vec{A} + \vec{A}\cdot\vec{\mathcal{P}}) + \left(\frac{e}{c}\right)^2 A^2\right] + e\Phi(\vec{r},t). \end{aligned}$$

Solving for the equations of motion of the displacement operator allows one to identify an oper-

ator that represents the velocity,

$$\begin{aligned}
\frac{d\vec{r}}{dt} &= \frac{[\vec{r}, H]}{i\hbar} \\
\frac{d}{dt} r_i &= \frac{1}{2mi\hbar} \{ r_i(\mathcal{P}_j - eA_j/c)(\mathcal{P}_j - eA_j/c) - (\mathcal{P}_j - eA_j/c)(\mathcal{P}_j - eA_j/c) r_i \} \\
&= \frac{1}{2mi\hbar} \{ r_i(\mathcal{P}_j - eA_j/c)(\mathcal{P}_j - eA_j/c) \\
&\quad - (\mathcal{P}_j - eA_j/c) r_i (\mathcal{P}_j - eA_j/c) + (\mathcal{P}_j - eA_j/c)(i\hbar\delta_{ij}) \} \\
&= \frac{(\mathcal{P}_i - eA_i/c)}{m} = \frac{\Pi_i}{m}, \\
\vec{\Pi} &\equiv \vec{\mathcal{P}} - e\vec{A}/c.
\end{aligned} \tag{3.5}$$

Here, $\vec{\Pi}$ is the canonical momentum, i.e. the mass multiplied by the velocity. It is the same combination of $\vec{\mathcal{P}}$ and $\vec{A}$ that appeared in kinetic energy term of the Hamiltonian in Eq. (3.4). The momentum $\vec{\mathcal{P}} = -i\hbar\nabla$ is NOT the operator associated with the mass multiplied by the velocity in the classical limit. Instead, the canonical momentum $\vec{\Pi}$ plays that role. In fact, the current density from the previous chapter is also redefined,

$$\vec{j}(\vec{r}, t) = \frac{-ie\hbar}{2m}[\psi^*(\vec{r}, t)\nabla\psi(\vec{r}, t) - (\nabla\psi^*(\vec{r}, t))\psi(\vec{r}, t)] - \frac{e^2\vec{A}}{mc}|\psi(\vec{r}, t)|^2. \tag{3.6}$$

Showing that this definition is consistent with the equation of continuity is left as a homework problem.

One can also solve for $d\vec{\Pi}/dt$, but first we calculate the commutator,

$$\begin{aligned}
[\Pi_i, \Pi_j] &= [(\vec{\mathcal{P}} - e\vec{A})_i, (\vec{\mathcal{P}} - e\vec{A}/c)_j] \\
&= (i\hbar)\left(e\partial_i A_j/c - e\partial_j A_i/c\right) = i\hbar e\epsilon_{ijk} B_k/c.
\end{aligned} \tag{3.7}$$

Moving forward,

$$\begin{aligned}
[\Pi_i^2, \Pi_j] &= \Pi_i[\Pi_i, \Pi_j] + [\Pi_i, \Pi_j]\Pi_i \\
\sum_i [\Pi_i^2, \Pi_j] &= (i\hbar e)\epsilon_{ijk}(B_k\Pi_i + \Pi_i B_k)/c \\
\frac{d}{dt}\vec{\Pi} &= \frac{i}{\hbar}[H, \vec{\Pi}] \\
&= \frac{e}{2mc}\left(\vec{\Pi} \times \vec{B} - \vec{B} \times \vec{\Pi}\right) - \nabla\Phi.
\end{aligned} \tag{3.8}$$

This becomes the usual, $d\vec{\mathcal{P}}/dt = e\vec{E} + e\vec{v} \times \vec{B}/c$, if one replaces $\vec{\mathcal{P}}$ with $\vec{\Pi}$ and if one ignores the fact that $\vec{B}$ and $\vec{\Pi} = m\vec{v}$ might not commute. They do not commute if $\vec{B}$ depends on $\vec{r}$ because $\vec{\mathcal{P}}$ has a gradient.

3.2 Gauge Invariance

Gauge invariance in electromagnetism shows that if one alters the vector and scalar potentials in the following manner,

$$\vec{A}(\vec{r},t) \to \vec{A}(\vec{r},t) + \nabla\Lambda(\vec{r},t), \quad \Phi(\vec{r},t) \to \Phi(\vec{r},t) - \frac{1}{c}\frac{\partial\Lambda(\vec{r},t)}{\partial t}, \tag{3.9}$$

that $\vec{E}$ and $\vec{B}$ are unchanged. However, the Hamiltonian is a function of $\vec{A}$ and Φ, not $\vec{E}$ and $\vec{B}$, which makes one question whether physics is invariant under a gauge transformation. Here, we show that even though the changes from the gauge transformation to $\vec{A}$ and Φ indeed yield a new Hamiltonian, the solutions are identical to those of the original Hamiltonian after applying a simple phase factor,

$$\psi(\vec{r},t) \to \exp\left[\frac{ie\Lambda(\vec{r},t)}{\hbar c}\right]\psi(\vec{r},t). \tag{3.10}$$

If one now considers the canonical momentum operator, altered by a gauge transformation, acting on a wave function, also altered by gauge invariance, one can see that the result is the same,

$$\begin{aligned}(-i\hbar\nabla - e\vec{A}c - e\nabla\Lambda(\vec{r},t)/c)\exp\left[\frac{ie\Lambda(\vec{r},t)}{\hbar c}\right]\psi(\vec{r},t) \\ = \exp\left[\frac{ie\Lambda(\vec{r},t)}{\hbar c}\right]\left[-i\hbar\nabla - e\vec{A}c - e\nabla\Lambda(\vec{r},t)/c - i\hbar\nabla\frac{ie\Lambda(\vec{r},t)}{\hbar c}\right]\psi(\vec{r},t) \\ = \exp\left[\frac{ie\Lambda(\vec{r},t)}{\hbar c}\right]\left[-i\hbar\nabla - e\vec{A}c\right]\psi(\vec{r},t).\end{aligned} \tag{3.11}$$

Following through with more factors of Π one can see that $\langle\phi|\Pi^n|\psi\rangle$ will be unchanged by the gauge transformation as long a both the operator Π and the states $|\psi\rangle$ and $\phi\rangle$ are transformed accordingly.

Similarly the term $i\hbar\partial_t - e\Phi$ acting on the altered wave function also shows an invariance.

$$\begin{aligned}\left[i\hbar\partial_t - e\Phi + \frac{1}{c}\partial_t\Lambda(\vec{r},t)\right]\exp\left[\frac{ie\Lambda(\vec{r},t)}{\hbar c}\right]\psi(\vec{r},t) \\ = \exp\left[\frac{ie\Lambda(\vec{r},t)}{\hbar c}\right]\left[i\hbar\partial_t - e\Phi(\vec{r},t)\right]\psi(\vec{r},t).\end{aligned} \tag{3.12}$$

Showing that the charge and current densities are unaffected by the transformation is left as a homework problem.

Aside: Checking Dimensions
Checking dimensions for relations involving electromagnetism can be painful. Here, the notation is consistent with cgs units in that $\vec{A}$ and Φ have the same dimension. Both $e\vec{A}$ and $e\Phi$ have dimensions of energy. Both $\vec{B}$ and $\vec{E}$ also have the same dimensions, as the dimensions of $e\vec{E}$ and $e\vec{B}$ are both energy per length. In these units, magnetic forces are given as $e\vec{v}\times\vec{B}/c$.

3.3 Interaction with Uniform Magnetic Fields

A constant magnetic field in the z direction can be described with the vector potentials,

$$A_z = 0,\ A_\rho = 0,\ A_\phi = \rho B/2, \tag{3.13}$$

which has azimuthal symmetry about the z axis as the vector potential winds around the z axis. Through a gauge transformation, one can also rewrite the vector potential as

$$A_y = Bx,\ A_x = 0, A_z = 0, \tag{3.14}$$

which gives the identical magnetic field. This appears to violate the azimuthal symmetry but has the advantage is that it is easy to solve. The wave function can be written in the form

$$\psi(x, y, z) = e^{ik_z z + ik_y y}\phi(x) \tag{3.15}$$

because both $\mathcal{P}_y$ and $\mathcal{P}_z$ commute with the Hamiltonian.

The differential equation for ϕ then turns out to be

$$\begin{aligned} E_{xy}\phi(x) &= -\frac{\hbar^2}{2m}\partial_x^2\phi(x) + \frac{1}{2m}(\hbar k_y - eBx/c)^2\phi(x) \\ &= -\frac{\hbar^2}{2m}\partial_x^2\phi(x) + \frac{e^2B^2}{2mc^2}(x - x_0)^2\phi(x), \\ x_0 &= \hbar c k_y/(eB), \\ E &= E_{xy} + \hbar^2 k_z^2/2m. \end{aligned} \tag{3.16}$$

This is the harmonic oscillator Hamiltonian with a spring constant $m\omega^2$ where

$$\omega = \frac{eB}{mc}\,. \tag{3.17}$$

The harmonic oscillator frequency is the same as the orbital frequency for a classical particle in a magnetic field of strength B, referred to as the cyclotron frequency.

The solutions thus look like a particle whose x position is centered about x_0, which is determined by k_y and whose solutions in the y and z directions are eigenstates of the momentum operators. It seems odd that a particle with circular motion would be an eigenstate of P_y. The solution to this paradox is that if P_y is constant, it does not imply that $v_y = (\hbar k_y - eA_y/c)/m$ is constant. Instead,

$$mv_y + eBx/c = \hbar k_y, \tag{3.18}$$

where p_y is the eigenvalue of $\mathcal{P}_y$. In fact this does describe circular motion centered about a point (x_c, y_c) with frequency $\omega = eB/mc$. To see this consider the general form for circular trajectory of frequency ω,

$$x = x_0 + R\cos(\omega t + \phi),\ \ y = y_0 + R\sin(\omega t + \phi), \tag{3.19}$$

whose velocities are

$$v_x = -\omega R\sin(\omega t + \phi),\ \ v_y = \omega R\cos(\omega t + \phi). \tag{3.20}$$

Inserting the expressions for x_0 and ω in Eq.s (3.16) and (3.17) into Eq.s (3.19) and (3.20), one sees that that Eq. (3.19) is simply a statement of circular motion with the x position centered around the minimum of the harmonic oscillator, which is determined by k_y. The center of the y motion is arbitrary, as this solution is a linear combination of orbits with different y_0 values.

The solutions are thus that of a harmonic oscillator for the motion in the $x-y$ plane multiplied by a plane-wave form for the z motion. The eigenenergies are

$$E_n(k_z) = \frac{\hbar^2 k_z^2}{2m} + (n+1/2)\hbar\omega, \tag{3.21}$$

where ω is the cyclotron frequency described above. Energies are independent of k_y. The energy levels are known as *Landau levels*. The degeneracy of the levels with respect to k_y will play the central role in describing the integral quantum Hall effect which we will discuss next semester.

3.4 Motion with Uniform Electric and Magnetic Fields

It is interesting to study the consequences of adding an electric field directed perpendicular to the magnetic field. Consider the case where the field is in the x direction. The Hamiltonian picks up an extra term, $-eEx$. In this case the differential equation for ϕ becomes

$$\begin{aligned} E_{xy}\phi(x) &= -\frac{\hbar^2}{2m}\partial_x^2\phi(x) + \frac{1}{2m}(\hbar k_y - eBx/c)^2\phi(x) - eEx\phi(x) \\ &= -\frac{\hbar^2}{2m}\partial_x^2\phi(x) + \frac{e^2B^2}{2mc^2}(x-x_0)^2\phi(x) - \frac{mc^2}{2}\left(\frac{E}{B}\right)^2\phi(x) - \frac{\hbar k_y E}{B}\phi(x), \\ x_0 &= \frac{\hbar c k_y}{eB} + \frac{mc^2 E}{eB^2}. \end{aligned} \tag{3.22}$$

The x position averaged over time would simply be x_0, the center of the harmonic oscillator. One can also calculate the average v_y,

$$\begin{aligned} v_y &= \frac{\Pi_y}{m} = \frac{\hbar k_y}{m} - \frac{eBx}{mc} \\ \bar{v}_y &= \frac{\hbar k_y}{m} - \frac{eBx_0}{mc} \\ &= -\frac{Ec}{B}. \end{aligned} \tag{3.23}$$

Here, the fact that the average x was x_0, as given in Eq. (3.22), was used to find the average velocity in the y direction. In the absence of an electric field, this average would be zero as expected given that classically, the motion is circular. Once the electric field is added, the particle moves, on average, in the $y-$direction with velocity $-(E/B)c$.

It might seem odd given that the average motion is in the $-\hat{y}$ direction while the electric field was applied in the $x-$direction. However, this is exactly the same result one finds for a classical trajectory. To gain an understanding of this, one can consider the motion in the absence of an electric field. This is clockwise circular motion of some radius R. Adding a small electric field in the x direction accelerates the particle on the upper part ($y > 0$) half of the trajectory, leading to

the speed being greatest when $x = +R$, and the speed being smallest when $x = -R$. Because the speed is higher when the particle is moving downward, and lower when the particle is moving upward, it is not surprising that the average v_y is negative. To understand why the particle does not drift in the direction of the electric field, one has to consider that the magnetic force acting on the particle when $x = +R$ is stronger than for $x = -R$ because the magnitude of the magnetic force is proportional to the speed. Thus the average magnetic force acting in the x direction is negative, which then allows it to cancel the added electric force.

One can also understand the result by considering the relativistic transformation properties of electromagnetic fields. If one begins with a magnetic field B_z in the z direction, then boosts in the y direction by a small velocity v_y, one generates an electric field in the x direction of strength $E_x = -v_y B_z$. Hence, if one has an electric field in the x direction, but views it in a frame moving with velocity $v_y = -E_x/B_z$, one sees no electric field and the motion is purely circular.

3.5 Path Integrals and the Aharanov-Bohm Effect

One way to perform Quantum Mechanics is through path integrals. Path integrals are usually a rather inconvenient way to go, but sometimes come in handy. The name refers to the fact that a sum over all intermediate states can be thought of as a path. To get a better idea we consider the matrix element

$$\begin{aligned}\langle\psi_f|e^{-iHt}|\psi_i\rangle &= \sum_{x_1\cdots x_{N-1}} \langle\psi_f|e^{-iH\delta t}|x_{N-1}\rangle\langle x_{N-1}|\cdots|x_2\rangle\langle x_2|e^{-iH\delta t}|x_1\rangle\langle x_1|e^{-iH\delta t}|\psi_i\rangle \\ &\approx \sum_{x_1\cdots x_{N-1}} \langle\psi_f|(1-iH\delta t)|x_{N-1}\rangle\langle x_{N-1}|\cdots|x_2\rangle\langle x_2|(1-iH\delta t)|x_1\rangle \\ &\qquad \langle x_1|(1-iH\delta t)|\psi_i\rangle.\end{aligned} \tag{3.24}$$

Here, $\delta t = t/N$ and the approximation becomes exact in the limit of large N.

If the set of states $x_1 \to x_n$ correspond to positions in coordinate space arranged in a mesh of size δx, the succesive points in the path, x_i and x_{i+1}, are constrained to be neighbors because H is local. One can appreciate the locality by returning to the picture from Chapter 2, where space was discretized in steps of δx and the kinetic term of the Hamiltonian only mixed neighboring sites,

$$\langle i|H|i+1\rangle = -\frac{\hbar^2}{2m\delta x^2}, \tag{3.25}$$

and the potential terms were diagonal. This means that each term in the sum can be thought of as a continuous trajectory where at each step in time the trajectory either remains at the same position or moves by $\pm\delta x$. This motivates the name *path integral*, though one might more accurately state that one sums over all trajectories rather than over all paths. The classical limit of quantum mechanics comes from the constraint of choosing the trajectory for which the phase becomes fixed with respect to small variations of the trajectory. This line of approximation is what motivates the simple calculations of interference phenomena, such as 2-slit interference, in elementary physics courses.

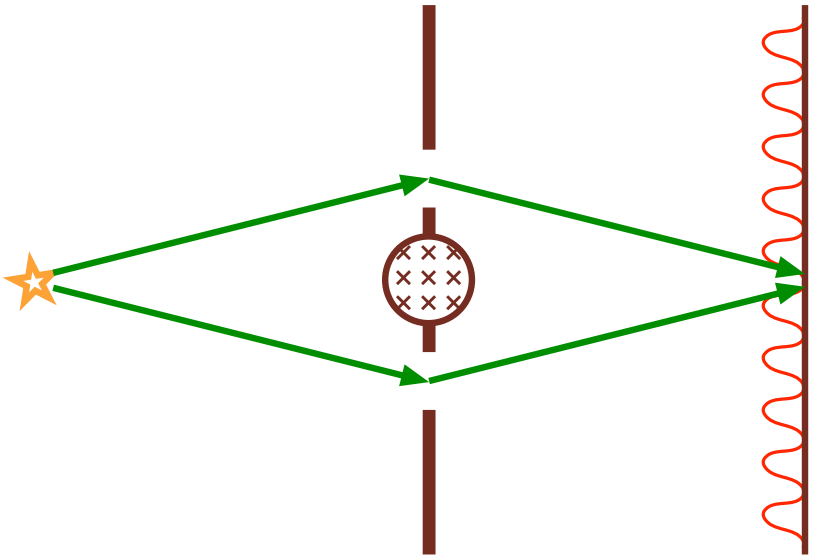


Figure 3.1: Illustration of the Aharanov-Bohm effect: The magnetic field $\vec{B}$ in the solenoid changes the relative phase of two paths of a charged particle which produce the interference pattern on the screen to the right. This modification derives from the changing the vector potential $\vec{A}$ in the region of the paths. The interference pattern shifts despite the fact that the paths never sample a region with non-zero $\vec{B}$.

In principle, all matrix elements can be considered in this fashion. Lattice gauge theory, which provides a powerful tool for numerically calculating the structure of the non-perturbative QCD vacuum where the trajectory accounts for all possible field configurations at each point in space-time, is built upon exactly such concepts. Path integral techniques are also often used in statistical mechanics by making the analogous decomposition,

$$e^{-\beta H} = \sum_{\alpha_1 \cdots \alpha_{N-1}} e^{-\delta\beta H}|\alpha_{N-1}\rangle\langle\alpha_{N-1}|\cdots|\alpha_2\rangle\langle\alpha_2|e^{-\delta\beta H}|\alpha_1\rangle\langle\alpha_1|e^{-\delta\beta H} \tag{3.26}$$

The interaction with an electromagnetic field $\vec{A}$ also contributes an off-diagonal piece due to the presence of the term $(-ie\hbar/mc)\vec{A}\cdot\nabla$. The off-diagonal term to the Hamiltonian is then

$$\langle i|H|i+1\rangle = -\frac{\hbar^2}{2m\delta x^2} - \frac{ie\hbar}{2mc\delta x}A_i = -\frac{\hbar^2}{2m\delta x^2}\left(1 + \frac{ie\delta x}{\hbar c}A_i\right). \tag{3.27}$$

Now, if we consider the evaluation of element $\langle\vec{r}_i|e^{-iHt}|\vec{r}_f\rangle$ as a sum over trajectories as described above. Each must include a series of links between neighboring sites. The effect of the vector potential is to add a product of the terms

$$\prod_{\text{links}}\left(1 + \frac{ie\delta x}{\hbar c}A_i\right) \tag{3.28}$$

to each trajectory.

Thus, the effect of the vector potential for propagation along a specific path from $\vec{r}_1$ to $\vec{r}_2$ is to modify the matrix element by a phase factor,

$$\langle\vec{r}_1|e^{-iHt}|\vec{r}_2\rangle \to \langle\vec{r}_1|e^{-iHt}|\vec{r}_2\rangle \exp\left[\frac{ie}{\hbar c}\int_{\text{path}} d\vec{r}\cdot\vec{A}\right], \tag{3.29}$$

where the path extends from $\vec{r}_1$ to $\vec{r}_2$. This added phase depends only on the path taken, but not the time-dependence of the trajectory because we have no time dependence and because the vector potential $\vec{A}$ is contracted only with $d\vec{r}$. Note that we have taken the liberty to jump to three dimensions at this point.

Now, we consider a two-slit interference experiment, where between the slits lies a small magnetic solenoid. In the upper path, the particle goes above the solenoid, while in the lower path the particle travels below the solenoid. Neither path samples the region inside the solenoid, where there is a magnetic field B and a magnetic flux, $\Phi_B = B\sigma$. Here, σ is the cross sectional area of the solenoid.

However, there is a vector potential outside the solenoid which must satisfy Stoke's theorem,

$$\oint \vec{A} \cdot d\ell = \Phi_B = |\vec{B}|\sigma. \tag{3.30}$$

The two path's contributions to the amplitude have phases which differ by this factor, multiplied by the charge,

$$\Delta\phi = \phi_{\text{upper path}} - \phi_{\text{lower path}} = \frac{e\Phi_B}{\hbar c}. \tag{3.31}$$

Here, the difference of the two paths is the same as adding the contribution to the phase around a closed loop, because subtracting the contribution from the lower path is the same as adding a contribution where one returns along the lower path. Because the expression for $\Delta\phi$ does not depend on the precise paths taken from the source to the screen, as long as one path is above the solenoid and one is below, it exactly sums up how any interference is adjusted between the upper and lower paths due to the presence of the vector potential $\vec{A}$. When the phase difference equals π, the maxima and minima of the interference pattern will have traded positions, and the interference pattern will have shifted by half a fringe. This phenomena has been observed experimentally with a thin magnetized iron filament called a *whisker*, as illustrated in Fig. 3.1.

Thus, altering the current through the solenoid shifts the interference pattern even though there are no magnetic or electric fields along the sampled trajectories. This makes it clear that $\vec{A}$ and Φ that are the fundamental fields in nature, not $\vec{E}$ and $\vec{B}$. Even though this means that $\vec{A}$ is fundamental, gauge transformations do not change any physical observables, even though a gauge transformation changes $\vec{A}$. This follows because the integral $\oint \vec{A} \cdot d\vec{\ell}$, is also independent of a gauge transformation, i.e. $\vec{A} \to \vec{A} + \nabla\Lambda$.

If one considers an eigenstate of a charged-particle's wave function, one can also consider the relative phases described by Eq. (3.29). The phase must return to zero, or a multiple of 2π if the phase is to have a single value return. For a charge q,

$$\frac{iq}{\hbar c}\int_{\text{path}} d\vec{r} \cdot \vec{A} = 2n\pi. \tag{3.32}$$

By Stoke's theorem this leads to a quantization of magnetic flux,

$$\Phi_B = \frac{2n\pi\hbar c}{q}. \tag{3.33}$$

Thus, the magnetic field in some fixed area cannot be assigned arbitrary values, but instead comes in integral units, if one is to have an eigenstate. Note that the two-slit experiment is not

an eigenstate. This seems peculiar because the integral units of flux depend on the charge. The Aharanov Bohm effect plays a critical role in Josephson's junctions. In that case the relevant charge is $2e$, the charge of a Cooper pair. This is not to state that one cannot have a non-integer amount of flux in a loop, only that the if the charged particles are in an eigenstate that they will affect the magnetic field in such a way that the flux in the loop settles into quantized levels. The quantized unit of flux is

$$\Phi_0 = \frac{2\pi\hbar c}{2e} = 2.068 \times 10^{-15}\ \mathrm{Tm}^2.$$

3.6 Exercises

1. Using the equations of motion for the wave function, show that the density and current defined by

$$\rho(\vec{r},t) = |\psi(\vec{r},t)|^2,$$
$$\vec{j}(\vec{r},t) = \frac{-i\hbar}{2m}(\psi^*(\vec{r},t)\nabla\psi(\vec{r},t) - (\nabla\psi^*(\vec{r},t))\psi(\vec{r},t)) - \frac{e\vec{A}}{mc}|\psi(\vec{r},t)|^2,$$

   satisfies the continuity equation,

$$\partial_t\rho + \nabla\cdot\vec{j} = 0.$$

2. Consider a particle of charge e traveling in the electromagnetic potentials

$$\mathbf{A}(\mathbf{r},t) = -\nabla\Lambda(\mathbf{r},t), \qquad \Phi(\mathbf{r},t) = \frac{1}{c}\frac{\partial\Lambda(\mathbf{r},t)}{\partial t}$$

   where $\Lambda(\mathbf{r},t)$ is an arbitrary scalar function.

   (a) What are the electromagnetic fields described by these potentials?
   (b) Show that the wave function of the particle is given by

$$\psi(\mathbf{r},t) = \exp\left[-\frac{ie}{\hbar c}\Lambda(\mathbf{r},t)\right]\psi_0(\mathbf{r},t),$$

   where ψ^0 solves the Schrödinger equation with $\Lambda = 0$.
   (c) Let $V(\mathbf{r},t) = e\Phi(t)$ be a spatially uniform time varying potential. Show that

$$\psi(\mathbf{r},t) = \exp\left[-\frac{ie}{\hbar}\int_{-\infty}^{t}\Phi(t')dt'\right]\psi_0(\mathbf{r},t)$$

   is a solution if ψ_0 is a solution with $\Phi = 0$.

3. For a gauge transformation, described in Eq. (3.9), including the associated the phase change to the wave function ψ, described in Eq. (3.10),

   (a) Show that the charge density $e\psi^*\psi$ is unchanged by the gauge transformation

(b) Show that the current

$$\vec{j} = \frac{1}{2m}\left[\psi^*(-i\hbar\nabla\psi) + (i\hbar\nabla\psi^*)\psi\right] - \frac{e}{mc}\vec{A}\psi^*\psi.$$

is unchanged.

(c) Show that $\langle\chi|H|\psi\rangle$ is unchanged in a gauge transformation where Λ is independent of time.

4. Find (or guess) the function $\Lambda(\vec{r},t)$ that corresponds to the gauge transformation in Eq. (3.9) responsible for re-expressing the vector potential in Eq. (3.13) to the form of Eq. (3.14), and show that both forms give the same magnetic field.

5. The expression for the $\bar{v}_y$ in Eq. (3.23) is only valid for non-relativistic velocities, where $|E| << |B|$. For a uniform magnetic field $B\hat{z}$, with no electric field, consider the form for the vector potential in Eq. (3.14). Performing a relativistic boost (Lorentz transformation), but for non-relativistic velocities, in the y direction by a velocity v_y, what is the resulting zero$^{\text{th}}$ component of the vector potential A_0? Equating this with the electric scalar potential, express the strength of the resulting electric field in terms of v_y and B.

6. In this problem, we reconsider the problem of a charged particle in the presence of both an electric and magnetic field, but do so in a different gauge. The electron is placed in a region of constant external magnetic field B directed along the z axis and of constant electric field E in the y direction.

(a) Choosing the vector potential to lie along the y axis and describe both the electric and magnetic fields, show that the Hamiltonian may be written in the form,

$$H = \frac{P_z^2}{2m} + \frac{P_x^2}{2m} + \frac{1}{2}m\omega^2(x - x_0 - v_0t)^2\ ,$$

and find ω, and v_0 in terms of E, B, e, m, k_y and c, where $\hbar k_y$ is the eigenvalue of P_y. Hint: Choose a gauge such that $\vec{E} = -(1/c)\partial_t\vec{A}$.

(b) Show that Schrödinger's equation, $i(\partial/\partial t)\Psi = H\Psi$ is satisfied by the form

$$\Psi(x,y,z,t) = e^{-i\epsilon_n t/\hbar + imv_0x/\hbar + ik_z z + ik_y y}\phi_n(x - x_0 - v_0t)\ ,$$

where ϕ_n refers to a harmonic-oscillator wave function characterized by the frequency ω and $\epsilon_n = (n + 1/2)\hbar\omega + mv_0^2/2$.

4 Angular Momentum and Central Potentials

Understanding the role of angular momentum in any quantum problem first requires taking stock of what rotational symmetries are either satisfied, or approximately satisfied. Because the understanding of symmetry is based on group theory, this chapter includes an extremely brief discussion of group theory before we launch into the various technical skills required to classify states based on angular momenta or to understand conservations and constraints of various transitions involving such symmetries.

4.1 The Baker-Campbell-Hausdorff Relation

Before our brief foray into group theory, we present the Baker-Camblel-Hausdorff relation, which is central to demonstrating the group nature of rotations. This relation expresses how two transformations might be expressed as a single transformation.

The Baker-Campbell-Hausdorff relation expresses how the product of two unitary transformations, e^{A} and e^{B}, are not necessarily the same as the transformation e^{A+B}. Instead, the relation is

$$e^{A}e^{B} = \exp\left\{A + B + \frac{1}{2}[A,B] + \frac{1}{12}[A,[A,B]] - \frac{1}{12}[B,[A,B]] + \cdots\right\}. \tag{4.1}$$

If the operators commuted, one would get the simple result $e^{A}e^{B} = e^{A+B}$. If when commuting A and B one gets an operator $C = [A,B]$ that commutes with both A and B, e.g. C is proportional to the unit matrix, the result is simpler

$$e^{A}e^{B} = e^{A+B}e^{C/2}. \tag{4.2}$$

For angular momentum $[L_i, L_j] = i\hbar\epsilon_{ijk}L_z$ and, because L_z does not commute with L_x and L_y, this simpler relation doesn't apply. Nonetheless, it helps clarify the role of commutators and is applicable in some other cases.

To demonstrate the relation where C is a constant, we consider the expansion

$$e^{A+B} = \sum_{N} \frac{(A+B)^n}{n!}. \tag{4.3}$$

Expanding $(A+B)^n$ gives all terms with all orderings of the n operators. For example, for $n = 5$ one of the terms is $ABBAB$. We wish to move all the A operators to the left, which requires commuting them past the B operators. Every time an A operator moves past a B operator one must add a term where the BA pair is replaced by $-C = [B,A]$. Using the binomial theorem, one can then write

$$\frac{(A+B)^n}{n!} = \sum_{i+j=n} \frac{A^iB^j}{n!}\frac{n!}{i!j!} + (-C)\frac{A^{i-1}B^{j-1}}{n!}\frac{n!}{i!j!}\bar{N}_1(i,j) + (-C)^2\frac{A^{i-2}B^{j-2}}{n!}\frac{n!}{i!j!}\bar{N}_2(i,j) \tag{4.4}$$

$$+ \cdots + (-C)^\ell\frac{A^{i-\ell}B^{j-\ell}}{n!}\frac{n!}{i!j!}\bar{N}_\ell(i,j) + \cdots$$

where $\bar{N}_\ell$ is the average number of ways to pick ℓ AB pairs from an order n term, under the constraint that the B operators in the pair were initially to the right of the A terms. This number is simply the number of such independent pairs times $(1/2)^\ell$ to account for the fact that only half the time does a given pair start off with the BA ordering,

$$\bar{N}_\ell = \left(\frac{1}{2}\right)^\ell \frac{i(i-1)\cdots(i-\ell+1)j(j-1)\cdots(j-\ell+1)}{\ell!}. \tag{4.5}$$

One can then factor the exponentials in the expression above to get Eq. (4.1),

$$\begin{aligned} e^{(A+B)} &= \sum_{ij\ell} \frac{A^i B^j}{i!j!}\frac{(-C/2)^\ell}{\ell!} \\ &= e^A e^B e^{-C/2}. \end{aligned} \tag{4.6}$$

Example 4.1: Rotations and Translations of Operators
One can rotate around the z axis by an angle α with the operator,

$$R(\alpha) = e^{\alpha\partial_\phi} = \sum_{n=0}^{\infty} \frac{1}{n!}(\alpha\partial_\phi)^n.$$

This demonstrates how the exponential form for rotations is related to a Taylor expansion. This can also be written as $e^{iL_z\phi/\hbar}$. Similarly, one can show how translations by a distance α are related to $e^{\alpha\partial_x} = e^{iP_x\alpha/\hbar}$. For an operator $B(\phi)$, show that with $R(\alpha) = e^{\alpha\partial_\phi}$ that

$$R(\alpha)B(\phi)R(-\alpha) = B(\phi+\alpha).$$

Solution:
First consider

$$e^{\alpha\partial_\phi}B(\phi) = \sum_n \frac{1}{n!}(\alpha^n\partial_\phi)^n B(\phi)\cdots$$

The additional $\cdots$ is to emphasize that there may be quantities further to the right that of $B(\phi)$ on which the derivatives act, i.e. there are functions of ϕ to the right of B. Next, one moves the derivatives from the left to the right. For each term ∂_ϕ^n, one must account for m occurrences where the derivative acts on $B(\phi)$ rather than on the quantities to the left.

$$e^{\alpha\partial_\phi}B(\phi) = \sum_{m=0}^{n}\sum_{n=0}^{\infty} \frac{\alpha^n}{n!}(\partial_\phi^n B(\phi))\partial_\phi^{n-m}\frac{n!}{m!(n-m)!}.$$

Here, the derivatives inside the parenthesis act only on $B(\phi)$, and given the substitution $\ell = m-n$,

$$\begin{aligned} e^{\alpha\partial_\phi}B(\phi) &= \left(\sum_m \frac{\alpha^m\partial_\phi^m}{m!}B(\phi)\right)\sum_\ell \frac{\alpha^\ell\partial_\phi^\ell}{\ell!} \\ &= \left(\sum_m \frac{(\alpha\partial_\phi)^m}{m!}B(\phi)\right)e^{\alpha\partial_\phi} \\ &= B(\phi+\alpha)e^{\alpha\partial_\phi}. \end{aligned}$$

Thus,

$$e^{\alpha\partial_\phi}B(\phi)e^{-\alpha\partial_\phi} = B(\phi+\alpha)e^{\alpha\partial_\phi}e^{-\alpha\partial_\phi} = B(\phi+\alpha).$$

To derive a more general relation, one could follow the same steps and show that for any operators A and B,

$$e^A B e^{-A} = B + [A,B] + \frac{1}{2}[A,[A,B]] + \frac{1}{3!}[A,[A,[A,B]]] + \cdots$$

4.2 Three-Dimensional Rotations

In quantum mechanics the angular momentum of a state, or an operator, tells us how the object behaves under rotation. Rotations are unitary transformations which can be classified as a *group*, and the operators that represent such transformations can be written in the form $e^{i\vec{L}\cdot\vec{\theta}/\hbar}$, where the $|\vec{\theta}|$ is the angle of the rotation about an axis the $\hat{\theta}$ direction. If an object does not depend on angle, then a rotation should have no effect on it. These are scalars, and examples are time, mass, or the dot product of two vectors, $\vec{u}\cdot\vec{v}$. Under rotations, both $\vec{u}$ and $\vec{v}$ change but the angle between the two vectors remains the same and $\vec{u}\cdot\vec{v}$ is unchanged.

The rotation group consists of unitary operators,

$$R(\vec{\alpha}) = e^{i\vec{L}\cdot\vec{\alpha}/\hbar}. \tag{4.7}$$

To understand why this a rotation, consider the case where $\vec{\alpha}$ is along the z axis. One can then see that

$$L_z = xP_y - yP_x = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) = -i\hbar\frac{\partial}{\partial\phi}, \tag{4.8}$$

which means that for a rotation of angle ϕ about the z-axis,

$$\begin{aligned} e^{iL_z\alpha/\hbar}f(\phi) &= e^{\alpha\partial_\phi} \\ &= \left(1 + \alpha\frac{\partial}{\partial\phi} + \frac{\alpha^2}{2!}\frac{\partial^2}{\partial\phi^2}\cdots\right)f(\phi) \\ &= f(\phi+\alpha). \end{aligned} \tag{4.9}$$

In order to be a true group, two consecutive rotations of $\vec{\alpha}$ and $\vec{\beta}$ must be identical to a single rotation $\vec{\gamma}$.

$$e^{i\vec{L}\cdot\vec{\beta}/\hbar}e^{i\vec{L}\cdot\vec{\alpha}/\hbar} = e^{i\vec{L}\cdot\vec{\gamma}/\hbar}. \tag{4.10}$$

Because the different components of $\vec{L}$ do not commute, it can be non-trivial to find the equivalent single rotation $\vec{\gamma}$ given $\vec{\alpha}$ and $\vec{\beta}$.

Given the definition $\vec{L} = \vec{r}\times\vec{p}$, it is straightforward to find the commutation relations,

$$[L_i, L_j] = i\hbar\epsilon_{ijk}L_k. \tag{4.11}$$

Our goal in this section is to discuss the requirement of using different operators, S_x, S_y and S_z to generate rotations, not in coordinate space but in a discrete vector space, meaning that $\vec{S}$ can

be expressed as matrices. The important requirement for $\boldsymbol{S}$ to be considered a rotation is that

$$e^{i\vec{S}\cdot\vec{\beta}/\hbar}e^{i\vec{S}\cdot\vec{\alpha}/\hbar} = e^{i\vec{S}\cdot\vec{\gamma}/\hbar}, \tag{4.12}$$

where the same $\vec{\gamma}$ results from a given $\vec{\alpha}$ and $\vec{\beta}$ as would have resulted if one had inserted $\vec{L} = -i\hbar\vec{r}\times\nabla$ instead of $\vec{S}$ everywhere above.

We wish to demonstrate that if the components of $\vec{S}$ obey the same commutation laws of the components of $\vec{L}$, the rotations will be identical. To see this we divide the two rotations into N smaller rotations with $N\to\infty$.

$$e^{i\vec{S}\cdot\vec{\beta}/\hbar}e^{i\vec{S}\cdot\vec{\alpha}/\hbar} = e^{i\vec{S}\cdot\vec{\beta}/(\hbar N)}e^{i\vec{S}\cdot\vec{\beta}/(\hbar N)}\cdots e^{i\vec{S}\cdot\vec{\beta}/(\hbar N)}e^{i\vec{S}\cdot\vec{\alpha}/(\hbar N)}e^{i\vec{S}\cdot\vec{\alpha}/(\hbar N)}\cdots e^{i\vec{S}\cdot\vec{\alpha}/(\hbar N)}. \tag{4.13}$$

To find the equivalent single rotation, one must expand each exponential then commute them in a manner as was done for the Baker-Campbell-Hausdorff lemma earlier. There are of order N^2 such commutations.

For our purposes we wish to consider the inner two exponentials,

$$\begin{aligned}
e^{i\vec{S}\cdot\vec{\beta}/(\hbar N)}e^{i\vec{S}\cdot\vec{\alpha}/(\hbar N)} &= 1 + i\vec{S}\cdot(\vec{\alpha}+\vec{\beta})/(\hbar N) \\
&\quad - \frac{1}{2(\hbar N)^2}\left(\vec{S}\cdot(\vec{\alpha}+\vec{\beta})\right)^2 + \frac{1}{2(\hbar N)^2}[\vec{S}\cdot\vec{\alpha},\vec{S}\cdot\vec{\beta}] + \mathcal{O}\frac{1}{(\hbar N)^3} \\
&= \exp\left\{i\left(\vec{S}\cdot(\vec{\alpha}+\vec{\beta})/(\hbar N) + i\frac{1}{2(\hbar N)^2}[\vec{S}\cdot\vec{\alpha},\vec{S}\cdot\vec{\beta}]\right)\right\} + \mathcal{O}\frac{1}{(\hbar N)^3} \\
&= \exp\left\{i\left(\vec{S}\cdot(\vec{\alpha}+\vec{\beta})/(\hbar N) - \frac{1}{2(\hbar N)^2}S_j\epsilon_{jk\ell}\alpha_k\beta_\ell\right)\right\} + \mathcal{O}\frac{1}{(\hbar N)^3} \\
&= \exp\left\{iS_j\left(\frac{\alpha_j+\beta_j}{(\hbar N)} - \frac{1}{2(\hbar N)^2}\epsilon_{jk\ell}\alpha_k\beta_\ell\right)\right\} + \mathcal{O}\frac{1}{(\hbar N)^3} \\
&= \exp\left\{i\frac{1}{\hbar}\vec{S}\cdot\delta\vec{\gamma}\right\} + \mathcal{O}\frac{1}{(\hbar N)^3} \\
\delta\vec{\gamma} &= \frac{\vec{\alpha}+\vec{\beta}}{N} - \frac{1}{2(\hbar N)^2}\vec{\alpha}\times\vec{\beta}.
\end{aligned} \tag{4.14}$$

Because there are of order N^2 such commutations we must perform to find $\vec{\gamma}$, we may throw away all terms of order $1/N^3$ or higher. The angle $\delta\gamma$ was determined purely from the commutation relations. Initially, there were $2N$ elements, but after this reduction there are $2N-1$ elements. One can then repeat the procedure and combine the element $e^{i\vec{S}\cdot\delta\gamma/\hbar}$ with one of the neighboring elements to reduce it further. Eventually, one can reduce everything to a single element described by a single angle $\vec{\gamma}$. Expressing $\vec{\gamma}$ in terms of $\vec{\alpha}$ and $\vec{\beta}$ is difficult, but that is not the point here. We only need to show that determining the angle depends purely on the commutation relations. Thus, if the commutation relations for the components for $\vec{S}$ are identical to the commutation relations for $\vec{L}$, then the equivalent angle $\vec{\gamma}$ will be the same in both cases.

The simplest example of three operators, S_x, S_y and S_z, which generate such rotations is the 2×2 representation,

$$S_x \equiv \frac{\hbar}{2}\sigma_x\,, \quad S_y \equiv \frac{\hbar}{2}\sigma_y\,, \quad S_z \equiv \frac{\hbar}{2}\sigma_z. \tag{4.15}$$

As a homework problem, you will able to find the equivalent angle $\vec{\gamma}$, describing two successive rotations $\vec{\alpha}$ and $\vec{\beta}$, by manipulating Pauli matrices. The arguments above also hold for any set of matrices where the commutation of any two such matrices results in something linearly proportional to the matrices. Although in such cases the mapping of $\vec{\alpha}$ and β to $\vec{\gamma}$ would differ from this case for three-dimensional rotations, the mapping would also be determined completely by the commutation relations.

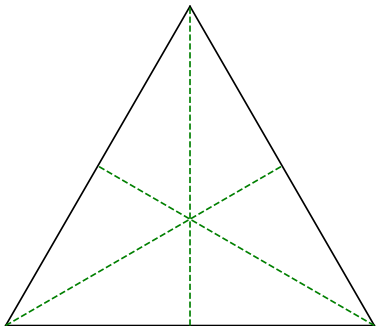
Many physical systems exhibit rotational symmetry, which leads to the fact that angular momentum is conserved. This can be seen by commuting the rotation operator for an infinitesimal rotation with the Hamiltonian. If $\boldsymbol{L_z}$ commutes with the Hamiltonian, then the eigenstates of $\boldsymbol{H}$ can be simultaneously chosen as eigenstates of $\boldsymbol{L_z}$. However, if the Hamiltonian has terms such as $\vec{\boldsymbol{L}} \cdot \vec{\boldsymbol{S}}$, neither $\boldsymbol{L_z}$ nor $\boldsymbol{S_z}$ commutes with the Hamiltonian. However, in that case the operator $\boldsymbol{J_z} \equiv \boldsymbol{L_z} + \boldsymbol{S_z}$ does commute. We will perform such an example in class later in this chapter.

4.3 A Greatly Abbreviated Guide to Group Theory

Groups are made of elements that represent transformations. The elements of a group are NOT the object being transformed, they are the transformations themselves. To be classified as a group, the group of elements $\boldsymbol{\mathcal{R}_i}$ must satisfy the following conditions:

1. The combination of any two elements returns an element of the group, $\boldsymbol{\mathcal{R}_i \mathcal{R}_j = \mathcal{R}_k}$. This is known as closure. Note this need not be commutative, $\boldsymbol{\mathcal{R}_i \mathcal{R}_j \neq \mathcal{R}_j \mathcal{R}_i}$.
2. Associativity, $\boldsymbol{(\mathcal{R}_i \mathcal{R}_j)\mathcal{R}_k = \mathcal{R}_i(\mathcal{R}_j \mathcal{R}_k)}$
3. There must exist an identity element $\mathbb{I}$, such that $\boldsymbol{\mathbb{I}\mathcal{R}_i = \mathcal{R}_i}$ and $\boldsymbol{\mathcal{R}_i \mathbb{I} = \mathcal{R}_i}$.
4. Every element $\boldsymbol{\mathcal{R}_i}$ must have an inverse element $\boldsymbol{\mathcal{R}_i^{-1}}$, such that $\boldsymbol{\mathcal{R}_i \mathcal{R}_i^{-1} = \mathbb{I}}$.

Often the elements of a group represent symmetry transformations. For example, consider the symmetries of an equilateral triangle, where the base lies along the $\boldsymbol{x}$ axis. The six group elements are:



The elements of the symmetry group of an equilateral triangle are the identity, two rotations, and three reflections about the dashed lines.

1. $\boldsymbol{\mathcal{R}_1}$, The identity
2. $\boldsymbol{\mathcal{R}_2}$, Rotating by 120°
3. $\boldsymbol{\mathcal{R}_3}$, Rotating by 240°
4. $\boldsymbol{\mathcal{R}_4}$, Reflecting about an axis through the center of the triangle in the $\boldsymbol{30°}$ direction
5. $\boldsymbol{\mathcal{R}_5}$, Reflecting about an axis through the center of the triangle in the $\boldsymbol{90°}$ direction
6. $\boldsymbol{\mathcal{R}_6}$, Reflecting about an axis through the center of the triangle in the $\boldsymbol{150°}$ direction

The group is defined by the integer coefficients $\boldsymbol{a(i,j) = 1 - 6}$, where $\boldsymbol{\mathcal{R}_i \mathcal{R}_j = \mathcal{R}_{a(i,j)}}$. The elements $\boldsymbol{a(i,j)}$ can be expressed as an $\boldsymbol{N \times N}$ matrix, known as a Cayley Table, https://en.wikipedia.org/wiki/Cayley_table.

Rotations are an example of a Matrix Lie Group, where in this context it simply means that there are a continuum of elements. For example, rotations can be taken about an axis in any direction, where the direction of the axis is described by the angles θ and ϕ in spherical coordinates, and the size of the rotation angle can be varied. Using $\hat{\alpha}$ to describe the rotation axis and $|\vec{\alpha}|$ to describe the rotation angle, the vector $\vec{\alpha}$ represents an element of the rotation group. The Lie groups used in physics are generally of the form

$$\mathcal{R}(\theta) = e^{\sum_{n=1,N} i\mathcal{G}_n\theta_n}. \tag{4.16}$$

The group elements are denoted by the $N-$dimensional vector θ, and the group elements are infinite in number. The matrices $\mathcal{K}_n$ are the generators of the group, and are often (but not always, e.g. the Lorentz transformation) Hermitian matrices, which makes each transformation unitary. As long as the commutation of any two generators gives a linear combination of generators, the multiplication of two group elements will be equivalent to a single element. These commutation rules are encapsulated by what are referred to as *structure constants* f_{ijk}

$$[\mathcal{G}_i, \mathcal{G}_j] = if_{ijk}\mathcal{G}_k. \tag{4.17}$$

The commutation rules for the N generators effectively define the entire group. Rotations are an example of a group, and for three dimensional rotations one needs three generators, corresponding to the rotations described by the three Euler angles.

To understand the role of the group conditions for the rotations of spin, one can imagine some quantum mechanical state with spins, and also perhaps with orbital angular momentum. If one rotates the entire system twice, by angles $\vec{\theta_1}$ and $\vec{\theta_2}$, where the direction is that of the rotational axis, it should be equivalent to a single rotation $\vec{\Theta}$. That single angle $\vec{\Theta}$ should not depend on whether one is rotating a spin wave function, or the orbital wave function. Even though rotating a spin-1/2 particle uses two-by-two matrices, and rotating a spin-one particle involve three-by-three matrices, the mapping of $\vec{\theta_1}$ and $\vec{\theta_2}$ to $\vec{\Theta}$ must be the same for both matrices. This will follow if the generators for the two cases have the same structure constants, even if the generators have different dimensionality, or in group theory parlance they are different representations of the same group but are isomorphic.

4.4 The $|\vec{L}|^2$ Operator

The angular momentum operator $L_z = -i\hbar\partial/\partial\phi$ commutes with the Hamiltonian if the Hamiltonian is invariant to rotations about the z axis, i.e. the Hamiltonian does not depend on ϕ. Furthermore, if the Hamiltonian is invariant to rotations about any axis, all three components of $\vec{L}$ commute with H. One may then define states which are simultaneously eigenstates of H and L_z, or L_x or L_z. But, one may not necessarily find states which are simultaneously eigenstates of L_x, L_y and L_z because these operators do not commute with one another.

However, the operator $L^2 = L_x^2 + L_y^2 + L_z^2$ is spherically symmetric and commutes with any of the three components of $\vec{L}$. This is because L^2 is a scalar, and scalars are not changed by rotation. One may therefore define eigenstates of a spherically symmetric Hamiltonian that are also eigenstates of both L^2 and L_z. We define the eigenvalues in terms of m and ℓ.

$$L_z|\ell, m\rangle = m\hbar|\ell, m\rangle \ , \quad L^2|\ell, m\rangle = \ell(\ell+1)\hbar^2|\ell, m\rangle. \tag{4.18}$$

The curious choice of $\ell(\ell+1)$ will become apparent in the next subsection, where it will be shown that this choice is consistent with ℓ being an integer.

Example 4.2: The $|\vec{L}|^2$ operator in Cartesian coordinates
Express $|\vec{L}|^2$ in terms of x, y, z and P_x, P_y, P_z, with all the coordinate space operators to the left of the momentum operators.
Solution:
One must be careful of the ordering of the spatial components r_i and the momentum operators P_j, because they do not commute. Given that $L_i = \epsilon_{ijk} r_j P_k$,

$$\begin{aligned}
|\vec{L}|^2 &= \epsilon_{ijk}\epsilon_{imn} r_j P_k r_m P_n \\
&= (\delta_{jm}\delta_{kn} - \delta_{jn}\delta_{mk}) r_j P_k r_m P_n \\
&= r_i P_j r_i P_j - r_i P_j r_j P_i \\
&= r^2 P^2 - i\hbar \vec{r}\cdot\vec{P} - r_i r_j P_i P_j + 3i\hbar\vec{r}\cdot\vec{P} \\
&= x^2(P_y^2 + P_z^2) + y^2(P_z^2 + P_x^2) + z^2(P_x^2 + P_y^2) \\
&- 2xyP_xP_y - 2yzP_yP_z - 2zxP_zP_x + 2i\hbar x P_x + 2i\hbar y P_y + 2i\hbar z P_z.
\end{aligned}$$

In classical physics the last three terms could be neglected because they are proportional to $\hbar$.

4.5 Raising and Lowering Operators for Angular Momentum

We define operators,

$$L_\pm \equiv L_x \pm i L_y. \tag{4.19}$$

These operators have the fortuitous property that

$$\begin{aligned}
[L_z, L_\pm] &= [L_z, L_x] \pm i[L_z, L_y] \\
&= i\hbar L_y \pm L_x \\
&= \pm\hbar L_\pm,
\end{aligned} \tag{4.20}$$

which means that

$$\begin{aligned}
L_z\left(L_\pm|m\rangle\right) &= L_\pm L_z|m\rangle \pm \hbar L_\pm|m\rangle \\
&= (m\pm 1)\hbar\left(L_\pm|m\rangle\right).
\end{aligned} \tag{4.21}$$

This means that $L_\pm$ effectively change an eigenstate of L_z to a new eigenstate with the eigenvalue either raised or lowered by $\hbar$.

One can also find the normalization of the new states by noting that

$$\begin{aligned}
\langle m|L_\mp L_\pm|m\rangle &= \langle m|L_x^2 + L_y^2 \pm i[L_x, L_y]|m\rangle \\
&= \langle m|L_x^2 + L_y^2 \mp \hbar L_z|m\rangle \\
&= \langle m|L^2 - L_z^2 \mp \hbar L_z|m\rangle \\
&= \left(\ell(\ell+1) - m^2 \mp m\right)\hbar^2.
\end{aligned} \tag{4.22}$$

Here, the eigenvalue of $|\vec{L}|^2$ is assigned the value of $\ell(\ell+1)$, which defines ℓ, but we have not yet shown that ℓ must be a integer or half integer.

In terms of normalized states, the raising and lowering operators acting on a state $|m\rangle$ become

$$L_\pm|m\rangle = \hbar[\ell(\ell+1) - m^2 \mp m]^{1/2}|m \pm 1\rangle. \tag{4.23}$$

By inspection, one sees that if the sequence of ms is to be finite that they must begin at $-\ell$ and end at ℓ, thus giving $2\ell+1$ values of m for a given ℓ. Because the ms must be separated by unit steps, ℓ and m must therefore be either integer or half integer.

4.6 Spherical Harmonics

After extensive application of the chain rule, the kinetic energy term in Schrödinger's wave equation may be written in spherical coordinates,

$$H_K = -\hbar^2\frac{\nabla^2}{2m} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right) - \frac{\hbar^2}{2mr^2}\left(\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\right). \tag{4.24}$$

Furthermore, the components of angular momentum may be written in terms of angular derivatives,

$$\begin{aligned}
L_z &= -ixP_y + iyP_x = -i\hbar\frac{\partial}{\partial\phi}, \\
L_x &= -iyP_z + izP_y = -i\hbar\left(-\sin\phi\frac{\partial}{\partial\theta} - \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right), \\
L_y &= -izP_x + ixP_z = -i\hbar\left(\cos\phi\frac{\partial}{\partial\theta} - \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right), \\
L_\pm &= -i\hbar e^{\pm i\phi}\left(\pm i\frac{\partial}{\partial\theta} - \cot\theta\frac{\partial}{\partial\phi}\right).
\end{aligned} \tag{4.25}$$

Because $L_z = -i\hbar\partial_\phi$, one can see that a Hamiltonian with rotational invariance about the z axis, i.e. no ϕ dependence, will commute with L_z. Thus, the conservation of L_z is tied to azimuthal symmetry. Further, the conservation of L_x or L_y depends on whether the Hamiltonian is invariant under rotations about the x or y axes.

Using the relation, $L^2 = L_z^2 + L_+L_- + i\hbar L_z$, one can see that the kinetic energy may be written as

$$\begin{aligned}
-\frac{\hbar^2}{2m}\nabla^2 &= -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right) + \frac{L^2}{2mr^2}, \\
|\vec{L}|^2 &= -\hbar^2\left\{-\frac{1}{\sin^2\theta}\left(\frac{\partial}{\partial\phi}\right)^2 - \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\right\}.
\end{aligned} \tag{4.26}$$

The last term reproduces the usual relation from classical physics for the effective centrifugal potential. If the potential is spherically symmetric, the Hamiltonian commutes with each component of L because they only involve angular derivatives, and each component of L commutes

with $\boldsymbol{L^2}$. Thus, we may write the solution as an eigenstate of a single component of $\vec{\boldsymbol{L}}$, such as $\boldsymbol{L_z}$, and also as an eigenstate of $\boldsymbol{L^2}$. However, one may not find a set of basis states that are simultaneously eigenstates of two different projections, e.g. $\boldsymbol{L_x}$ and $\boldsymbol{L_z}$, because the three projections do not commute amongst themselves.

If the potential $\boldsymbol{V(\vec{r}) = V(r)}$ is a function only of $\boldsymbol{r}$, and is independent of both $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$, $\boldsymbol{L_z}$ and $\boldsymbol{|\vec{L}|^2}$ commute with the Hamiltonian and the eigenstates of $\boldsymbol{H}$ are also eigenstates of $\boldsymbol{L_z}$ and $\boldsymbol{L^2}$. These eigenstates will be labeled by $\boldsymbol{m}$ and $\boldsymbol{\ell}$,

$$\boldsymbol{\Psi(\vec{r}) = \phi_\ell(r) Y_{\ell,m}(\theta,\phi).} \tag{4.27}$$

Being an eigenstate of $\boldsymbol{L^2}$, the operator $\boldsymbol{L^2}$ in Eq. (4.26) can be replaced with its eigenvalue $\boldsymbol{\hbar^2\ell(\ell+1)}$. The spherical harmonic then factors out, and there is no mention of $\boldsymbol{m}$ in Schrödinger's equation, which justifies having $\boldsymbol{\phi_\ell(r)}$ being independent of $\boldsymbol{m}$. Schrödinger's equation for an energy $\boldsymbol{E}$ is then

$$\boldsymbol{E\phi_\ell(r) = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right)\phi_\ell(r) + \frac{\hbar^2\ell(\ell+1)}{2mr^2}\phi_\ell(r) + V(r)\phi_\ell(r).} \tag{4.28}$$

The radial function has a label $\boldsymbol{\ell}$ because the radial wave function, Eq. (4.28), depends on $\boldsymbol{\ell}$, but not on $\boldsymbol{m}$.

Our immediate goal is to understand the angular functions $\boldsymbol{Y_{\ell,m}(\theta,\phi)}$ which are eigenstates of $\boldsymbol{L_z}$ and $\boldsymbol{L^2}$ and are referred to as spherical harmonics. In terms of bras and kets,

$$\boldsymbol{Y_{\ell,m}(\theta,\phi) = \langle\hat{n}|\ell,m\rangle,} \tag{4.29}$$

which implies the normalization,

$$\boldsymbol{\int d\phi d\cos\theta\, |Y_{\ell,m}(\theta,\phi)|^2 = 1.} \tag{4.30}$$

The eigenstates of $\boldsymbol{L_z}$ must go as $\boldsymbol{e^{im\phi}}$. Combined with the requirement that the raising operator in Eq. (4.25) working on $\boldsymbol{Y_{\ell,\ell}}$ gives zero, one can derive the expression for $\boldsymbol{Y_{\ell,\ell}(\theta,\phi)}$,

$$\begin{aligned}
\boldsymbol{Y_{\ell,\ell}} &= \boldsymbol{f_\ell(\theta)e^{i\ell\phi},} \qquad (4.31)\\
\boldsymbol{\left(\pm i\frac{\partial}{\partial\theta} - \cot\theta\frac{\partial}{\partial\phi}\right) f_\ell(\theta)e^{i\ell\phi}} &= \boldsymbol{0,}\\
\boldsymbol{(\partial_\theta - \ell\cot\theta)\, f_\ell(\theta)} &= \boldsymbol{0,}\\
\boldsymbol{f_\ell(\theta)} &= \boldsymbol{c_\ell \sin^\ell(\theta),}\\
\boldsymbol{Y_{\ell,\ell}(\theta)} &= \boldsymbol{c_\ell e^{i\ell\phi}\sin^\ell(\theta),}
\end{aligned}$$

where the normalization is given by

$$\boldsymbol{c_\ell = \left[\frac{(-1)^\ell}{2^\ell\ell!}\right]\sqrt{\frac{(2\ell+1)(2\ell)!}{4\pi}}.} \tag{4.32}$$

The factor $\boldsymbol{(-1)^\ell}$ above is a matter of convention.

By operating on the known $\boldsymbol{Y_{\ell,m}}$s with $\boldsymbol{L_-}$ one may generate $\boldsymbol{Y_{\ell,m}}$s for successively lower values of $\boldsymbol{m}$. Because the $\boldsymbol{Y_{\ell,m}}$s are eigenstates of $\boldsymbol{L_z}$ the $\boldsymbol{m}$ dependence is always trivial as it goes proportional to $\boldsymbol{e^{im\phi}}$, but the $\boldsymbol{\theta}$ dependence can be messy. Examples of a few spherical harmonics are

$$\begin{aligned} \boldsymbol{Y_{0,0}} &= \frac{\boldsymbol{1}}{\boldsymbol{\sqrt{4\pi}}} \\ \boldsymbol{Y_{1,0}} &= \boldsymbol{\sqrt{\frac{3}{4\pi}}\cos\theta} \\ \boldsymbol{Y_{1,\pm1}} &= \boldsymbol{\mp\sqrt{\frac{3}{8\pi}}\sin\theta e^{\pm i\phi}} \\ \boldsymbol{Y_{2,0}} &= \boldsymbol{\sqrt{\frac{5}{16\pi}}(3\cos^2\theta - 1)} \\ \boldsymbol{Y_{2,\pm1}} &= \boldsymbol{\mp\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta e^{\pm i\phi}} \\ \boldsymbol{Y_{2,\pm2}} &= \boldsymbol{\sqrt{\frac{15}{32\pi}}\sin^2\theta e^{\pm 2i\phi}} \end{aligned} \tag{4.33}$$

One might ask why half integer values of $\boldsymbol{\ell}$ are never mentioned for spherical harmonics. The problem is that the wave functions then become discontinuous as $\boldsymbol{\phi}$ goes past $\boldsymbol{2\pi}$. Thus, half-integral angular momenta can only be used for intrinsic spins and not as labels for a spatial wave function because the spatial part of the Hamiltonian involves spatial gradients.

Finally, Legendre polynomials are defined as

$$\boldsymbol{P_\ell(\cos\theta) \equiv \sqrt{\frac{4\pi}{2\ell+1}}Y_{\ell,m=0}(\theta).} \tag{4.34}$$

The $\boldsymbol{Y_{\ell,m}}$s with odd $\boldsymbol{\ell}$ have odd parity, i.e. under the transformation $(\boldsymbol{\theta \to \pi - \theta,\ \phi \to \phi + \pi})$ the $\boldsymbol{Y_{\ell,m}}$s with odd $\boldsymbol{\ell}$ switch sign. This will play an important role in determining that many matrix elements will be zero. Note that the radial wave function, $\boldsymbol{\phi_\ell(r)}$, is always even under parity.

4.7 Wigner $\mathcal{D}$ Matrices

This is for the most point an exercise in notation, but it does emphasize that rotations mix states within a given multiplet of angular momentum $\boldsymbol{\ell}$, where the degeneracy is $\boldsymbol{2\ell+1}$. Rather than expressing rotations as a function of $\boldsymbol{\alpha_x}$, $\boldsymbol{\alpha_y}$ and $\boldsymbol{\alpha_z}$, one can express a rotation as a function of the three Euler angles, which represent consecutive rotations about the $\boldsymbol{z}$, $\boldsymbol{y}$ and the new $\boldsymbol{z}$ axes.

$$\boldsymbol{\mathcal{D}^\ell_{mm'}(\alpha,\beta,\gamma) = \exp\left(\frac{-iS_z\alpha}{\hbar}\right)\exp\left(\frac{-iS_y\beta}{\hbar}\right)\exp\left(\frac{-iS_z\gamma}{\hbar}\right).} \tag{4.35}$$

The label $\boldsymbol{\ell}$ refers to the dimension of the matrices, $\boldsymbol{2\ell+1}$, used to represent the rotations, e.g. for two component matrices $\boldsymbol{\ell = 1/2}$, while $\boldsymbol{m}$ and $\boldsymbol{m'}$ refer to the components of the matrices, $\boldsymbol{-\ell \le m \le \ell}$.

For instance, for $\ell = 1/2$,

$$\begin{aligned}\mathcal{D}^{1/2} &= \exp\left(\frac{-i\sigma_z\alpha}{2}\right)\exp\left(\frac{-i\sigma_y\beta}{2}\right)\exp\left(\frac{-i\sigma_z\gamma}{2}\right) \\ &= [\cos(\alpha/2) - i\sigma_z\sin(\alpha/2)][\cos(\beta/2) - i\sigma_y\sin(\beta/2)][\cos(\gamma/2) - i\sigma_z\sin(\gamma/2)] \\ &= \begin{pmatrix} e^{-i(\alpha+\gamma)/2}\cos(\beta/2) & -e^{-i(\alpha-\gamma)/2}\sin(\beta/2) \\ e^{i(\alpha-\gamma)/2}\sin(\beta/2) & e^{i(\alpha+\gamma)/2}\cos(\beta/2) \end{pmatrix}.\end{aligned} \tag{4.36}$$

Example 4.3: Rotations within a Spin$-1/2$ Multiplet Using Wigner $\mathcal{D}$ matrices

Using Wigner $\mathcal{D}$ matrices, rotate the state,

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

about the x axis by an angle ψ.
Solution:
For Euler angles, the first rotation is about the z axis, followed by a rotation about the y axis, and the last rotation is about the new z axis. In order to rotate about the x axis, one must first rotate the system with about the z axis by angle $\alpha = -\pi/2$, then rotate about the y axis by ψ. Finally, to restore the coordinate system, you need to rotate again about the z axis by $\gamma = \pi/2$. Using Eq. (4.36), the rotation matrix is

$$\mathcal{D}^{1/2}(\pi/2, \psi, -\pi/2) = \begin{pmatrix} \cos(\psi/2) & -e^{-i\pi/2}\sin(\psi/2) \\ e^{i\pi/2}\sin(\psi/2) & \cos(\psi/2) \end{pmatrix}.$$

One can check that this is the rotation one would expect, $e^{-i\vec{\sigma}\cdot\vec{\psi}/2}$,

$$\begin{aligned} e^{-i\vec{\sigma}\cdot\vec{\psi}/2} &= \cos\psi\mathbb{I} - i\sigma_x\sin\psi \\ &= \begin{pmatrix} \cos(\psi/2) & -i\sin(\psi/2) \\ -i\sin(\psi/2) & \cos(\psi/2) \end{pmatrix}.\end{aligned}$$

Then, multiplying $|\uparrow\rangle$,

$$\mathcal{D}^{1/2}(\pi/2, \psi, 0)|\uparrow\rangle = \begin{pmatrix} \cos(\psi/2) \\ -i\sin(\psi/2) \end{pmatrix}.$$

By inspection, one can see that if one rotates by π the state will become $|\downarrow\rangle$ as expected. Further, if the rotation is by $\psi = \pi/2$ the resulting state will be an eigenstate of σ_y with eigenvalue -1. Again, this would be expected for a rotation about the x axis.

The $\mathcal{D}$ matrices can also rotate operators, if the operators carry the labels of angular momentum, e.g J and M. For example, if one considers the operators X, Y and Z, one can express three

similar operators

$$\begin{aligned} \boldsymbol{R_0} &= \boldsymbol{Z}, \\ \boldsymbol{R_+} &= (\boldsymbol{X} + i\boldsymbol{Y})/\sqrt{2}, \\ \boldsymbol{R_-} &= -(\boldsymbol{X} - i\boldsymbol{Y})/\sqrt{2}. \end{aligned} \tag{4.37}$$

By inspecting the form for spherical harmonics, one can see that these operators rotate just like the spherical harmonics, $\boldsymbol{Y_{\ell=1,m}}$. Thus, even though they are operators, not states, their rotations can be expressed with the same $\boldsymbol{\mathcal{D}^{\ell=1}}$ matrices used to rotate states with labels $\boldsymbol{\ell = 1}$ and $\boldsymbol{m}$. This may seem uneasy, as we are more accustomed to seeing operators, $\boldsymbol{A}$, transforming as $\boldsymbol{UAU^\dagger}$. However, in this case the operators do not have two indices $\boldsymbol{m, m'}$, i.e. $\boldsymbol{A_{mm'}}$, and instead have a single index $\boldsymbol{m}$. This realization plays a critical role in understanding the Wigner Eckart theorem later in the course.

4.8 Separating Relative and Center-of-Mass Coordinates

Consider the Schrödinger equation for two particles interacting through a potential $\boldsymbol{V(\vec{r}_1 - \vec{r}_2)}$. The kinetic energy term,

$$\boldsymbol{H_K = -\frac{\hbar^2}{2m_1}\nabla_1^2 - \frac{\hbar^2}{2m_2}\nabla_2^2} \tag{4.38}$$

needs to be rewritten such that the derivates are with respect to center-of-mass and relative coordinates,

$$\boldsymbol{\vec{R} \equiv \frac{m_1\vec{r}_1 + m_2\vec{r}_2}{m_1 + m_2}, \;\; \vec{r} \equiv \vec{r}_1 - \vec{r}_2.} \tag{4.39}$$

Using these definitions, one can show that that the kinetic energy becomes

$$\boldsymbol{H_K = -\frac{\hbar^2}{2M}\nabla_r^2 - \frac{\hbar^2}{2\mu}\nabla_r^2,} \tag{4.40}$$

where $\boldsymbol{M \equiv m_1 + m_2}$ and $\boldsymbol{\mu \equiv m_1 m_2/(m_1 + m_2)}$.

If the potential is only a function of $\boldsymbol{r}$, with no mention of $\boldsymbol{R}$, the total momentum operator, $\boldsymbol{-i\hbar\nabla_R}$, commutes with the Hamiltonian. The wave function may then be written as a product of center-of-mass and relative coordinates,

$$\boldsymbol{\Psi(\vec{r}, \vec{R}) = e^{iK\cdot\vec{R}}\phi_{\text{rel.}}(\vec{r}),} \tag{4.41}$$

with the center-of-mass dependence being characterized by a plane wave. The overall energy is then a sum of the eigenenergy of the relative wave function plus $\boldsymbol{\hbar^2 K^2/2M}$.

Note that if one of the masses is much larger than the other that the reduced mass $\boldsymbol{\mu}$ approaches the smaller of the two masses. If both masses are equal, the reduced mass is half the mass of either of the two individual masses. For our purposes, we will solve problems such as the hydrogen atom assuming the potential is fixed. For the real case, one need only replace the mass with the reduced mass to include the effect that the source of the potential is itself mobile.

Separating the center-of-mass coordinates is convenient whenever one has a potential that is a function of $\vec{r}_1 - \vec{r}_2$ only. Writing Schrödinger's equation,

$$\left\{-\frac{\hbar^2}{2M}\nabla_R^2 - \frac{\hbar^2}{2\mu}\nabla_r^2\right\} e^{iK\cdot\vec{R}}\phi(\vec{r}) + V(r)e^{iK\cdot\vec{R}}\phi(\vec{r}) = Ee^{iK\cdot\vec{R}}\phi(\vec{r}), \tag{4.42}$$

One can factor out $e^{iK\cdot\vec{R}}$ after operating with ∇_R^2 and get a simple equation for the relative wave function $\phi(\vec{r})$.

$$-\frac{\hbar^2}{2\mu}\nabla_r^2\phi(\vec{r}) + V(r)\phi(\vec{r}) = E_{\rm rel}\phi(\vec{r}) \tag{4.43}$$

$$E = \frac{\hbar^2K^2}{2M} + E_{\rm rel}$$

The factorization above works for any non-relativistic problem whenever the potential is a function of $\vec{r}$ and not $\vec{R}$.

Example 4.4: Two Masses Coupled by a Spring
One especially easy example of factorization is the three-dimensional harmonic oscillator in Cartesian coordinates. In this case the problem is even more factorizable. Again, using $\vec{R} = (\vec{r}_1 + \vec{r}_2)/2$ and $\vec{r} = \vec{r}_1 - \vec{r}_2$,

$$H = -\frac{\hbar^2\nabla_R^2}{2M} + \frac{\hbar^2\nabla_r^2}{2\mu} + V(\vec{r}), \tag{4.44}$$

$$V(\vec{r}) = \frac{1}{2}k|\vec{r}|^2 = \frac{1}{2}k\left[x^2 + y^2 + z^2\right].$$

This potential is spherically symmetric, and is thus a candidate to solve using wave functions that are products of radial wave functions and $Y_{\ell m}$s. However, because this potential can be written as the sum of an x-dependent, a y-dependent and a z-dependent piece, it also works to factorize the wave function as

$$\phi(\vec{R},\vec{r}) = \phi_x(x)\phi_y(y)\phi_z(z)e^{i\vec{K}\cdot\vec{R}}, \tag{4.45}$$

where x, y and z are the components of $\vec{r}$. One can separate the Schrödinger equation for the three-dimensional relative wave function into three one-dimensional Schrödinger equations.

$$-\frac{\hbar^2}{2\mu}\partial_x^2\phi_x + \frac{1}{2}kx^2\phi_x(x) = E_x\phi_x(x) \tag{4.46}$$

$$-\frac{\hbar^2}{2\mu}\partial_y^2\phi_y + \frac{1}{2}ky^2\phi_y(y) = E_y\phi_y(y)$$

$$-\frac{\hbar^2}{2\mu}\partial_z^2\phi_z + \frac{1}{2}kz^2\phi_z(z) = E_z\phi_z(z)$$

Here, $E = E_x + E_y + E_z + \hbar^2K^2/2M$. By multiplying the first equation by $\phi_y\phi_z$, the second equation by $\phi_x\phi_z$ and the third by $\phi_x\phi_y$, then adding the three equations, one finds that they provide a solution to the three-dimensional equation.

Thus, one is able to take a six-dimensional equation, factor out the center-of-mass motion, and recognize the factorizability, and reduce the problem to three trivial one-dimensional equations of motion. The energies are

$$\boldsymbol{E_x = (n_x + 1/2)\hbar\omega},\ \boldsymbol{E_y = (n_y + 1/2)\hbar\omega},\ \boldsymbol{E_z = (n_z + 1/2)\hbar\omega}, \tag{4.47}$$

where $\boldsymbol{\omega = \sqrt{k/\mu}}$. The total energy is then the sum

$$\boldsymbol{E = \frac{\hbar^2 K^2}{2m} + (n_x + n_y + n_z + 3/2)\hbar\omega}. \tag{4.48}$$

We refer to these solutions as the Cartesian-basis solutions to the harmonic oscillator. Because the harmonic oscillator presented above has spherical symmetry, one can also write solutions in spherical coordinates where each solution has labels $\boldsymbol{\ell}$ and $\boldsymbol{m}$ and has a factor $\boldsymbol{Y_{\ell,m}(\theta,\phi)}$. The spherical basis will be discussed in Sec. (4.10).

4.9 Coupled Oscillators in Cartesian Coordinates

One convenient aspect of harmonic oscillator potentials is that they can be solved for the $\boldsymbol{N}$-body case even when $\boldsymbol{N > 2}$. Here, we consider $\boldsymbol{N}$ particles of mass $\boldsymbol{m}$ interacting through mutual harmonic oscillator potentials, with every mass being coupled to every other mass. Consider the potential,

$$\boldsymbol{V(\vec{r}_1, \cdots \vec{r}_N) = \frac{1}{2}k\left(\sum_{i<j}|\vec{r}_i - \vec{r}_j|^2\right)}. \tag{4.49}$$

This potential is independent of the center-of-mass coordinates. That is, a translation of all coordinates does not affect the potential. One can therefore write the solution as a product of a center-of-mass wave function and a wave function that depends on $\boldsymbol{N-1}$ relative coordinates.

The above problem would be intractable if not for a trick which is unique to the harmonic oscillator. One adds a fictitious potential that depends only on the center-of-mass coordinate.

$$\boldsymbol{V_f = \frac{1}{2}k|\vec{r}_1 + \vec{r}_2 + \cdots \vec{r}_N|^2 = \frac{1}{2}N^2 k|R_{cm}|^2}. \tag{4.50}$$

With this choice, the sum of the real and fictitious potentials cancels all the cross terms and becomes,

$$\boldsymbol{V_{tot} = V_f + V = \frac{1}{2}Nk\left(r_1^2 + r_2^2 + \cdots r_N^2\right)}. \tag{4.51}$$

The total energy can be written as the sum of the center-of-mass energy plus the relative energy,

$$\boldsymbol{E_{tot} = (n_f + 3/2)\hbar\sqrt{Nk/m} + E_{rel}}, \tag{4.52}$$

where $\boldsymbol{E_{rel}}$ is the energy of relative motion which is our ultimate goal. Here the term $\boldsymbol{(n_f + 3/2)\hbar\sqrt{Nk/m}}$ replaces the usual kinetic energy of the center-of-mass. The single factor of

N inside the square root comes from the two factors of N in the expression for the fictitious potential canceled by the one factor of N in the total mass,

$$\omega_f = \sqrt{\frac{N^2 k}{M}}, \quad M = Nm. \tag{4.53}$$

The total energies are those of N harmonic oscillators with mass m and spring constant Nk, and have eigenenergies

$$E_{\text{tot}} = (n_1 + n_2 + \cdots n_N + 3N/2)\hbar\sqrt{Nk/m}. \tag{4.54}$$

The ground state energy we are interested in is thus $E_{\text{tot}}(n_i = 0)$ minus the energy of the center of mass, $(3/2)\hbar\sqrt{Nk/m}$.

$$\begin{aligned} E_{0,\text{rel}} &= \frac{3N}{2}\hbar\sqrt{Nk/m} - \frac{3}{2}\hbar\sqrt{Nk/m} \\ &= \frac{3(N-1)}{2}\sqrt{Nk/m}. \end{aligned} \tag{4.55}$$

For $N = 2$, the ground state energy is that of a single oscillator with reduced mass of $\mu = m/2$. The wave function of the ground state can be written as the product of all the ground state wave functions, divided by the wave function of the center-of-mass. If one wanted to include the center-of-mass motion, one could assign a momentum to the center of mass, $\vec{P}$, then add $|\vec{P}|^2/2m$ to the energies, and add a factor of $e^{i\vec{P}\cdot\vec{R}_{\text{cm}}/\hbar}$ to the wave function.

Example 4.5: Three Coupled Harmonic Oscillators
Consider three identical masses m, each coupled to the other by a potential $k|\vec{r}_i - \vec{r}_j|^2/2$.

1. What are the energies of the ground state and the first excited state?
2. What is the degeneracy of the first excited state?

Solution:
The ground state energy is found by applying Eq. (4.55) with $N = 3$.

$$E_0 = \frac{3(3-1)}{2}\sqrt{3k/m} = \frac{3}{2}\sqrt{6k/m}.$$

The first excited state energy is found by taking the first excited state energy from the total energy in Eq. (4.54),

$$E_{\text{tot}} = (3/2 + n_{\text{tot}})\hbar\sqrt{Nk/m},$$

where n_{tot} is some non-negative integer. Then, one must subtract the energy of the center of mass, noting that the center-of-mass might also be excited, i.e.,

$$E_{\text{c.m.}} = (3/2 + n_f)\hbar\sqrt{Nk/m}.$$

The difference gives the energy

$$E_{\text{rel}} = (3/2 + n_{\text{tot}} - n_f)\hbar\sqrt{Nk/m}.$$

We are only looking for the excitation energy for given state of the center of mass, so we put the center of mass in its ground stat, $n_f = 0$, and see that the first excited state of the relative energy is higher by an amount $\hbar\sqrt{Nk/m}$, with $N = 3$.
Finally, finding the degeneracy is a bit tricky. There exist $3N$ possibilities of $n_1 + n_2 + n_3 + n_N = 1$. The factor of three accounts for the fact that these are three-dimensional harmonic oscillators. However, this includes the possibility that the center of mass was excited. Thus, the answer is $3(N-1)$. So, for three coupled oscillators, the degeneracy of the first excited state is 6.

4.10 Solving the Radial Wave Equation for Spherically Symmetric Cases

Similar factorization ideas are applied any time one deals with a spherically symmetric potential. In that case one can write the wave function as a product of $\phi_\ell(r)$ and $Y_{\ell,m}(\theta,\phi)$ and reduce the problem to a one-dimensional problem of the radial coordinate. The one-dimensional Schrödinger equation for $\phi_\ell(r)$ then becomes

$$-\frac{\hbar^2}{2\mu}\left(\partial_r^2 + \frac{2}{r}\partial_r\right)\phi_\ell(r) + \left(\frac{\hbar^2\ell(\ell+1)}{2\mu r^2} + V(r)\right)\phi_\ell(r) = E\phi_\ell(r). \tag{4.56}$$

Although this is a one-dimensional differential equation, it is not a one-dimensional Schrödinger equation due to the extra derivative term $(2/r)\partial_r$. The one-dimensional Schrödinger form can be regained by defining

$$u_\ell(r) \equiv r\phi_\ell(r). \tag{4.57}$$

The wave equation for u looks like a 1-d Schrödinger equation with a centrifugal potential,

$$-\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial r^2}u(r) + \left(\frac{\hbar^2\ell(\ell+1)}{2\mu r^2} + V(r)\right)u(r) = Eu(r). \tag{4.58}$$

The boundary condition for u is that it must go to zero at the origin so that ϕ is finite, and analytic, at the origin. The basic strategy is to solve the equation for $u(r)$, then divide by r to get $\phi_\ell(r)$.

The wave equation is particularly simple for s waves ($\ell = 0$) because such solutions reduce to simple one-dimensional problems with an infinite potential for $r < 0$. Considering non-zero ℓ introduces a divergent centrifugal potential at the origin, proportional to $\ell(\ell+1)/r^2$. In the neighborhood of $r = 0$ the solution looks like either

$$u_\ell(r) \sim r^{\ell+1}\left(1 + \mathcal{O}r + \mathcal{O}r^2\cdots\right), \text{the regular solution}, \tag{4.59}$$

or

$$u_\ell(r) \sim r^{-\ell}\left(1 + \mathcal{O}r + \mathcal{O}r^2\cdots\right), \text{the irregular solution}. \tag{4.60}$$

Equivalently, one could state that the solutions for $\phi_\ell = u_\ell/r$ behave as r^ℓ and $r^{-\ell-1}$ for the regular and irregular solutions respectively. Clearly, only the regular solution satisfies the boundary conditions. Even when a Coulomb potential is added, which is also divergent at the origin only less so, the behavior at the origin can be expanded as shown above. If the divergence of the

potential at the origin is as strong or stronger than $1/r^2$, one needs to rethink the expansion above.

When the potential $V(r)$ is zero, i.e. free particles, the regular solutions are known as spherical bessel functions, $j_\ell(kr)$, while the irregular solution are referred to as spherical Neumann functions, $n_\ell(kr)$. These correspond to solutions to the wave equations for ϕ_ℓ above, rather than for u_ℓ. Before writing these down, we rewrite Schrödinger's equation with $V = 0$ in terms of $x = kr$ with $k = \sqrt{2mE/\hbar^2}$. Multiplying Schrödinger's equation by $2\mu/(\hbar^2 k^2)$,

$$-\left(\partial_x^2 + \frac{2}{x}\partial_x\right)\phi_\ell(x) + \frac{\ell(\ell+1)}{x^2}\phi_\ell(x) = \phi_\ell(x). \tag{4.61}$$

For the first few ℓs, the solutions are

$$\begin{aligned}
j_0(x) &= \frac{\sin x}{x}, & n_0(x) &= -\frac{\cos(x)}{x}, \\
j_1(x) &= \frac{\sin x}{x^2} - \frac{\cos x}{x}, & n_1(x) &= -\frac{\cos x}{x^2} - \frac{\sin x}{x}, \\
j_2(x) &= \left(\frac{3}{x^3} - \frac{1}{x}\right)\sin x - \frac{3}{x^2}\cos x, & n_2(x) &= -\left(\frac{3}{x^3} - \frac{1}{x}\right)\cos x - \frac{3}{x^2}\sin x, \\
&\vdots & &\vdots
\end{aligned} \tag{4.62}$$

By taking a linear combination of j_ℓ and n_ℓ, one can find a solution which behaves like an outgoing wave at large r,

$$\begin{aligned}
h_\ell(kr) &= j_\ell(kr) + i n_\ell(kr) \\
&\approx (-i)^{\ell+1}\frac{e^{ikr}}{kr}\ , \text{ as } r \to \infty.
\end{aligned} \tag{4.63}$$

These are known as Hankel functions. As an example, for $\ell = 1$,

$$h_1(x) = \frac{e^{ix}}{x}\left(-1 - \frac{i}{x}\right). \tag{4.64}$$

Example 4.6: Spherical Wells

1. Solve for the lowest energies of the $\ell = 0$ and $\ell = 1$ states of an infinite spherical well of radius R. Note that for the $\ell = 1$ solutions, a transcendental expression will remain.

   Solution: The BC and solutions for $\ell = 0$ are

$$\begin{aligned}
j_0(kR) &= 0, \\
\sin(kR) &= 0, \\
k &= \frac{\pi}{R}, \\
E_0 &= \frac{\hbar^2\pi^2}{2mR^2}.
\end{aligned}$$

For $\ell = 1$,

$$\begin{aligned} j_1(kR) &= 0, \\ \frac{\sin(kR)}{kR} - \cos(kR) &= 0, \\ \tan(kR) &= kR, \ (\text{transcendental eq.for } k) \\ E_1 &= \frac{\hbar^2 k^2}{2m}. \end{aligned}$$

The values of kR for which $j_\ell(kR) = 0$ are known as the zeros of the spherical Bessel function. The values are:

ℓ	$n=1$	$n=2$	$n=3$	$\cdots$
0	π	2π	3π	$\cdots$
1	4.4934	7.72525	**10.9041**	$\cdots$
2	5.7635	9.0950	12.3299	$\cdots$
3	6.9879	10.4171	13.6980	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Values of the n^{th} zeros of spherical Bessel function of order ℓ

From the table, one can see that the first excited state has $\ell = 1$, and that the first $\ell = 2$ state is also of lower energy than the second $\ell = 0$ state.

2. Outline how one would solve for the bound state of a well of finite depth $-V_0$ and width R for $\ell = 1$.

   Solution: One would choose $j_1(kr)$ for the interior solution (region I), with

   $$k = \sqrt{2m(V_0 - B)/\hbar^2},$$

   then match to a solution for negative energy (imaginary k) outside. The outer solution would be chosen as

   $$Ah_1(iqr) = -A\frac{e^{-qr}}{iqr}\left(1 + \frac{1}{qr}\right),$$

where $q = \sqrt{2mB/\hbar^2}$ and A is an unknown constant. One then needs to satisfy the two BC (continuous value and slope) at $r = R$ by varying the binding energy B and the constant A. This would lead to a transcendental equation for B.

4.11 Spherical Harmonic Oscillator – Spherical Basis

The spherically symmetric harmonic oscillator can be approached either through Cartesian coordinates as performed earlier or in a spherical basis. First, we review the Cartesian solutions.

Because the potential can be written as

$$V(\vec{r}) = \frac{1}{2}kr^2 = \frac{1}{2}k\left(x^2 + y^2 + z^2\right), \tag{4.65}$$

the solutions factorize into x, y and z-dependent functions,

$$\Psi(\vec{r}) = \psi_{n_x}(x)\psi_{n_y}(y)\psi_{n_z}(z). \tag{4.66}$$

Each piece is a solution of the 1-dimensional Schrödinger's equation and the total energy is

$$E = (N + 3/2)\hbar\omega\,, \;\; N \equiv n_x + n_y + n_z,\; \omega\sqrt{k/\mu}. \tag{4.67}$$

The $N = 0$ and $N = 1$ eigenstates have the form

$$\begin{aligned}
\phi_{n_x=n_y=n_z=0}(\vec{r}) &\sim e^{-r^2/(2a^2)} \\
\phi_{n_x=1,n_y=n_z=0}(\vec{r}) &\sim xe^{-r^2/(2a^2)} \\
\phi_{n_x=0,n_y=1,n_z=0}(\vec{r}) &\sim ye^{-r^2/(2a^2)} \\
\phi_{n_x=n_y=0,n_z=1}(\vec{r}) &\sim ze^{-r^2/(2a^2)}
\end{aligned} \tag{4.68}$$

By looking at the form of the $Y_{\ell,m}$s, one can see that the $N = 0$ state has a $\ell = 0, m = 0$ angular dependence, and can therefore be written as a spherically symmetric function, $e^{-r^2/(2a_0^2)}$, multiplied by $Y_{0,0}$. The solution for the $n_x = n_y = 0, n_z = 1$ state can be written as a product of $Y_{1,0}$ and the radial function $re^{-r^2/(2a_0^2)}$. By taking linear combinations, $\phi_{n_x=1,n_y=n_z=0} \pm \phi_{n_x=0,n_y1,n_z=0}$, one finds solutions which can be written as the same radial wave function multiplied by $Y_{1,\pm1}$.

In spherical coordinates the labels n_x, n_y and n_z are replaced by N, ℓ and m. Mapping the solutions for higher N is a bit tricky. For $N \geq 2$, one can count the states from a Cartesian perspective. One needs to know the number of ways to get three integers to add to N. First, the number of ways, $d_\perp$, to get two positive integers to add to $N_\perp$ is

$$d_\perp(N_\perp) = \sum_{n_x=0,N_\perp} = N_\perp + 1. \tag{4.69}$$

Requiring that a third integer adds to N gives the total degeneracy $d(N)$,

$$d(N) = \sum_{N_\perp=0,N} d_\perp(N_\perp) = \frac{(N+1)(N+2)}{2}. \tag{4.70}$$

Thus, there is one way to get $N = 0$, 3 ways to get $N = 1$, six ways to get $N = 2$, etc.

To determine which ℓ multiplets combine to create the d_N Cartesian solutions with a given energy, $(N + 3/2)\hbar\omega$, consider two pieces of evidence: First, creating the Cartesian state with $n_z = 0, n_x = n_y = 0$ is represented by the spherical state $Y_{\ell=N,m=N}$, therefore there must

be at least one $\ell = N$ multiplet. Secondly, all the multiplets for $N - 2$ must have corresponding multiplets which are generated by operating on those states with the spherically symmetric operator $a_x^2 + a_y^2 + a_z^2$. We can now determine the multiplets by:

$$d(N) = \frac{(N+1)(N+2)}{2} = (2N+1) + d(N-2) + \text{ any others.} \tag{4.71}$$

Solving for the number of *others*, one finds there are no others. Hence the excitation $N + 2$ states have the same states, but with one more multiplet of $\ell = N + 2$. Because one knows that for $N = 0$ there is one $\ell = 0$ state and for $N = 1$ there is one $\ell = 1$ multiplet, one can quickly find all the multiplets for any N.

As an example the $N = 5$ states are covered by one $\ell = 5$ multiplet, one $\ell = 3$ multiplet and one $\ell = 1$ multiplet. Note that all states with even N have even parity and all states with odd N have odd parity.

4.12 The Hydrogen Atom

There are three standard problems of spherically symmetric potentials where the solutions are analytic, the inifinite well, the harmonic oscillator and the Coulomb potential. Here, we consider the case where the potential is attractive,

$$V(r) = -\frac{e^2}{r}. \tag{4.72}$$

One may rewrite the Schrödinger equation,

$$\left(-\frac{\partial^2}{\partial r^2} + \frac{\ell(\ell+1)}{r^2} - \frac{Z_1 Z_2}{a_0 r}\right) u_\ell(r) = -k^2 u_\ell(r), \tag{4.73}$$

where the Bohr radius is defined $a_0 \equiv \hbar^2/(\mu e^2)$ and $k^2 = -2\mu E/\hbar^2$. For large r the potential and centrifugal terms are negligible and the wave function must behave as $e^{\pm ikr}$ multiplied by terms that vary more slowly in r.

The solutions to the Schrödinger equation can be written in terms of associated Laguerre polynomials.

$$R_{n,\ell}(r) = \frac{u_{n,\ell}}{r} = \left\{\left(\frac{2}{na_0}\right)^3 \frac{(n-\ell-1)!}{2n[(n+\ell)!]^3}\right\}^{1/2} e^{-r/(na_0)} \left(\frac{2r}{na_0}\right)^\ell L_{n+\ell}^{2\ell+1}\left(\frac{2r}{na_0}\right). \tag{4.74}$$

For a given ℓ there are many solutions labeled by the integer $n > \ell$. The eigenenergies can be written simply as

$$E_n = -\frac{e^2}{2a_0}\frac{1}{n^2}. \tag{4.75}$$

If larger charges are used, the above expressions are modified by scaling a_0 by $1/(Z_1 Z_2)$.

Writing a few solutions,

$$\begin{aligned} R_{1,0} &= \frac{2}{a_0^{3/2}} e^{-r/a_0} \\ R_{2,0} &= \frac{1}{(2a_0)^{3/2}} \left(2 - \frac{r}{a_0}\right) e^{-r/(2a_0)} \\ R_{2,1} &= \frac{1}{(2a_0)^{3/2}} \frac{r}{a_0\sqrt{3}} e^{-r/(2a_0)} \end{aligned} \tag{4.76}$$

The degeneracy, where energies with different ℓ have the same energies seems accidental, but is related to the similar degeneracy in the harmonic oscillator. Both degeneracies can be explained by considering the Lenz vector, which commutes with the Hamiltonian for the special case of the Coulomb potential.

$$\vec{Q} \equiv \frac{1}{2\mu}\left(\vec{p} \times \boldsymbol{L} - \boldsymbol{L} \times \vec{p}\right) - \frac{e^2}{r}\vec{r}. \tag{4.77}$$

This operator is Hermitian and, if the commutation between $\vec{p}$ and $\vec{L}$ is ignored, is identical to the classical expression for the Lenz vector.

One can also define a scaled operator,

$$\vec{K} \equiv \sqrt{\frac{-m}{2H}}\vec{Q}, \tag{4.78}$$

which is a little odd because one is defining the square root of an operator. However, because we are considering only eigenstates of the Hamiltonian with negative energies, this is not too sick.

One can show that the components of $\vec{K}$ and the components of the angular momentum $\vec{L}$ obey simple commutation relations,

$$[K_i, K_j] = i\hbar\epsilon_{ijk}L_k\ , \quad [K_i, L_j] = i\hbar\epsilon_{ijk}K_k \tag{4.79}$$

These commutation relations are reminiscent of angular momentum commutation relations, and in fact, if one defines two new operators,

$$\begin{aligned} \vec{M} &\equiv \frac{\vec{L} + \vec{K}}{2} \\ \vec{N} &\equiv \frac{\vec{L} - \vec{K}}{2}, \end{aligned} \tag{4.80}$$

one can see that $\mathcal{M}$ and $\mathcal{N}$ obey the same commutation relations as L,

$$[M_i, M_j] = i\hbar\epsilon_{ijk}M_k, \quad [N_i, N_j] = i\hbar\epsilon_{ijk}N_k, \quad [M_i, N_j] = 0 \tag{4.81}$$

Because $\vec{M}$ and $\vec{N}$ are linear combination of the $\vec{L}$ and $\vec{K}$, they commute with the Hamiltonian, and because they commute with one another they may simultaneously be defined.

Furthermore, some algebra reveals that the Hamiltonian may be written as

$$H = -\frac{me^4}{2(\vec{K}^2 + \vec{L}^2 + \hbar^2)} = \frac{me^4}{2(2\vec{M}^2 + 2\vec{N}^2 + \hbar^2)} \tag{4.82}$$

Finally, note that the Lenz vector is always perpendicular to the angular momentum, which means that $\vec{r} \cdot \vec{L} = \vec{K} \cdot \vec{L} = 0$. From the definitions of $\vec{M}$ and $\vec{N}$, one then sees the constraint that

$$M = N \tag{4.83}$$

where M and N are the quantum numbers denoting the magnitude of the vectors $\vec{M}$ and $\vec{N}$ in the same way that ℓ denotes the magnitude of $\vec{L}$. The eigenenergies are thus,

$$E_M = -\frac{me^4}{2\hbar^2(4M(M+1)+1)} = -\frac{me^4}{2\hbar^2(2M+1)^2} \tag{4.84}$$

Because M could be either integral over half integral, the numbers

$$n \equiv 2M + 1 \tag{4.85}$$

are integral values.

These states are not eigenstates of L, but given the fact that there must exist eigenstates of L, one can conclude that several states of the same ℓ might be degenerate.

4.13 Adding Angular Momentum

In a spherically symmetric potential, the orbital angular momentum, L, commutes with the Hamiltonian. Thus, one can choose eigenstates of L_z and $|\vec{L}|^2$ as a subset of states, knowing that such states can also be eigenstates of the Hamiltonian. The eigenvalues of are $\hbar m$ and $\hbar^2\ell(\ell+1)$, the labels ℓ and m are often called "good" quantum numbers, because they form a basis that is convenient for diagonalizing the Hamiltonian. Many particles also have intrinsic spin, even those particles which are currently considered as point particles such as electrons and photons. Thus, in addition to the orbital quantum numbers two more quantum numbers may be used to describe the eigenstates of a single particle in the potential, s and m_s, which describe the magnitude and projection of the spin angular momentum.

Often a term exists in the Hamiltonian which couples the two types of spin. One example is the spin-orbit interaction,

$$H_{\text{s.o.}} = \alpha \vec{L} \cdot \vec{S}. \tag{4.86}$$

This term originates from relativistic considerations which we will see later in the course. Because the term is written as a rotational scalar and does not involve an external field, which would explicitly break the rotational symmetry, we expect that the overall angular momentum remains conserved. Indeed, one can see that each component of the total angular momentum

$$\vec{J} \equiv \vec{L} + \vec{S} \tag{4.87}$$

commutes with the spin-orbit term, even though none of the components of $\vec{L}$ or $\vec{S}$ commute with $H_{\text{s.o.}}$ individually. Furthermore, the total squared orbital and total spin angular momentum, L^2 and S^2, also commute with $H_{\text{s.o.}}$. Thus, there are two new quantum numbers j and m_j which replace m_ℓ and m_s as good quantum numbers, while m_ℓ and m_s are no longer good quantum numbers.

A clearer insight into the spin-orbit term can be attained by rewriting it explicitly in terms of j, ℓ and s.

$$\left(\vec{L}+\vec{S}\right)^2 = \vec{L}^2 + \vec{S}^2 + 2\vec{L}\cdot\vec{S} \tag{4.88}$$
$$\vec{L}\cdot\vec{S} = \frac{1}{2}\left(\vec{J}^2 - \vec{L}^2 - \vec{S}^2\right),$$

which means that the spin-orbit term for a state in the basis labeled by j, m_j, ℓ and s may be expressed as,

$$H_{\text{s.o.}} = \frac{\alpha\hbar^2}{2}\left[j(j+1) - \ell(\ell+1) - s(s+1)\right]. \tag{4.89}$$

The coupling of spins is a common occurrence in all branches of physics. In nuclear physics, the spin-orbit term is surprisingly large, and is responsible for the basic scheme for nuclear shell structure. In describing hadron spectroscopy, a spin-spin interaction is largely responsible for the difference of the spin $3/2$ delta baryon and the spin $1/2$ proton which are comprised of quarks of the same flavor. The coupling of angular momentum in physics thus often involves changing from the ℓ, s, m_ℓ, m_s basis to the ℓ, s, j, m_j basis.

When undergoing a change of basis, the number of states involved is $(2s+1)(2\ell+1)$ as can be determined by considering the number of combinations of m_ℓ and m_s. However, changing to the j, m_j basis only mixes states with identical $m_j = m_\ell + m_s$. Because the states in a multiplet described by j must be complete, running from $-j$ to j, we see that the maximum value of j is

$$j \leq \ell + s. \tag{4.90}$$

Because there is only one state with $m_\ell = \ell$ and $m_s = s$, there is only one state with $m_j = \ell + s$ and thus only one j multiplet with $j = \ell + s$. Counting the number of pairs of m_ℓ and m_s that add up to a specific value of m_j, and realizing that every j multiplet must be complete lets one see that the values of j involved are

$$j_{\max} = \ell + s \tag{4.91}$$
$$j_{\min} = |\ell - s|.$$

These are known as the triangle relations, as they can be considered as constraints involved in adding vectors. One cannot add two vectors of lengths ℓ and s and obtain a vector of length j outside this range. A common error made by students is to only consider the $j = \ell + s$ and $j = \ell - s$ multiplets, and forget the multiplets with intermediate values of j.

One can check that the net number of states is unchanged. Each j multiplet has $2j+1$ states. Assuming $s \leq \ell$ the number of j multiplets is $(2s+1)$ while the average j of the multiplets is ℓ, which means that the number of states is $\langle 2j+1\rangle(2s+1) = (2\ell+1)(2s+1)$. This is consistent with the total number of states calculated in the m_ℓ, m_s basis.

4.14 Clebsch-Gordan Coefficients

Changing from the m_ℓ, m_s basis to the j, m_j basis is described by the overlap of matrix elements,

$$\langle \ell, s, j, m_j | \ell, s, m_\ell, m_s\rangle. \tag{4.92}$$

Such matrix elements are known as Clebsch-Gordan coefficients and are referred to through a variety of confusing notations, such as $C(\ell, s, m_\ell, m_s; j, m)$ and nearly every other possible permutation of the arguments. Sometimes the coefficients are labeled by superscripts and subscripts and sometimes they look like matrix elements $\langle j, m_j | \ell, s, m_\ell, m_s \rangle$, where ℓ and s are implied in the bra. The notations are remarkably confusing given that the only purpose of not writing down eight labels, with two repeated as in Eq. (4.92), is that the labels ℓ and s, which are always the same in the bra and ket, are not written down twice. To emphasize that the coefficients behave as matrix elements we will refer to the coefficients either in the form $\langle j, m_j | \ell, s, m_\ell, m_s \rangle$ or with ℓ and s repeated in the bra as in Eq. (4.92).

The matrix elements are usually used as part of a basis transformation, e.g. expressing states labeled by $|\ell, s, j, m_j\rangle$ as a linear combination of $|\ell, s, m_\ell, m_s\rangle$. Using completeness,

$$\sum_{m_\ell, m_s} |\ell, s, m_\ell, m_s\rangle\langle \ell, s, m_\ell, m_s| = \mathbb{I}, \tag{4.93}$$

$$\begin{aligned} |\ell, s, j, m_j\rangle &= \sum_{m_\ell, m_s} |\ell, s, m_\ell, m_s\rangle\langle \ell, s, m_\ell, m_s|\ell, s, j, m_j\rangle \\ &= \sum_{m_\ell, m_s} \rangle\langle \ell, s, m_\ell, m_s|\ell, s, j, m_j\rangle|\ell, s, m_\ell, m_s\rangle \\ &= \sum_{m_\ell, m_s} \langle j, m_j|\ell, s, m_\ell, m_s\rangle|\ell, s, m_\ell, m_s\rangle. \end{aligned}$$

The last line was justified by the fact that the Clebsch-Gordan coefficients are real (because the raising and lowering operators don't introduce any complex phases).

Finding the matrix elements is straight-forward, though tedious. They can be found using the algebra for raising and lowering angular momentum. First, remember that the matrix elements are all proportional to $\delta_{m_j, m_\ell + m_s}$. Because there is only one multiplet with $m_j = \ell + s$ that matrix element is simple to express.

$$\langle \ell, s, j = \ell + s, m_j = \ell + s | \ell, s, m_\ell = \ell, m_s = s\rangle = 1. \tag{4.94}$$

To generate the coefficients involving the same $j = \ell + s$ but reduced m_j, one can use the lowering operators,

$$\begin{aligned} |\ell, s, j, m_j - 1\rangle &= \frac{1}{\sqrt{j(j+1) - m_j(m_j - 1)}} J_- |\ell, s, j, m_j\rangle \\ &= \frac{1}{\sqrt{j(j+1) - m_j(m_j - 1)}} \left(L_- + S_-\right) |\ell, s, j, m_j\rangle. \end{aligned} \tag{4.95}$$

Applying this to the case where $m_j = j = \ell + s$ one generates an expression for the matrix

elements with $m_j = \ell + s - 1$,

$$\begin{aligned}|\ell, s, j = \ell + s, m_j = \ell + s - 1\rangle &= \frac{1}{\sqrt{j(j+1) - m_j(m_j-1)}} \\ &\quad \cdot (L_- + S_-)\,|\ell, s, m_\ell = \ell, m_s = s\rangle \\ &= \frac{1}{\sqrt{j(j+1) - m_j(m_j-1)}} \\ &\quad \cdot \Big(\sqrt{\ell(\ell+1) - m_\ell(m_\ell-1)}|\ell, s, m_\ell = \ell - 1, m_s = s\rangle \\ &\quad + \sqrt{s(s+1) - m_s(m_s-1)}|\ell, s, m_\ell = \ell, m_s = s - 1\rangle\Big).\end{aligned} \tag{4.96}$$

One can now read off the Clebsch-Gordan coefficients. For instance,

$$\langle \ell, s, j = \ell + s, m_j = \ell + s - 1|\ell, s, m_\ell = \ell - 1, m_s = s\rangle = \frac{\sqrt{\ell(\ell+1) - m_\ell(m_\ell-1)}}{\sqrt{j(j+1) - m_j(m_j-1)}}. \tag{4.97}$$

Finding the Clebsch-Gordan coefficient for $j \neq \ell + s$ is a bit trickier. By knowing that the $j = \ell + s - 1$ states are orthogonal to the $j = \ell + s$ states, allows one to express the $j = \ell + s - 1$ states down by inspection. For instance,

$$\begin{aligned}|\ell, s, j = \ell + s - 1, m_j = \ell + s - 1\rangle = \\ \frac{1}{\sqrt{j(j+1) - m_j(m_j-1)}} \cdot \Big(\sqrt{\ell(\ell+1) - m_\ell(m_\ell-1)}|\ell, s, m_\ell = \ell, m_s = s - 1\rangle \\ - \sqrt{s(s+1) - m_s(m_s-1)}|\ell, s, m_\ell = \ell - 1, m_s = s\rangle\Big)\end{aligned} \tag{4.98}$$

Of course, one could multiply the states by any arbitrary phase and the coefficients would work as well. The convention is that the coefficient $\langle j_1, j_2, j, j|j_1, j_2, m_1 = j_1, m_2 = j - j_1\rangle$ is real and positive.

Example 4.7: Calculating Clebsch-Gordan Coefficients
Find $\langle j = 3/2, m = 3/2|j_1 = 3/2, j_2 = 1, m_1 = 3/2, m_2 = 0\rangle$.
First, using $J_- = J_{1,-} + J_{2,-}$

$$\begin{aligned}|j = 5/2, m = 5/2\rangle &= |m_1 = 3/2, m_2 = 1\rangle \\ \sqrt{(5/2)(7/2) - (5/2)(3/2)}|j = 5/2, m = 3/2\rangle &= \\ \sqrt{(3/2)(5/2) - (3/2)(1/2)}|m_1 = 1/2, m_2 = 1\rangle &\\ + \sqrt{(1)(2) - (1)(0)}|m_1 = 3/2, m_2 = 0\rangle, &\\ |j = 5/2, m = 3/2\rangle = &\\ \sqrt{\frac{3}{5}}|m_1 = 1/2, m_2 = 1\rangle + \sqrt{\frac{2}{5}}|m_1 = 3/2, m_2 = 0\rangle. &\end{aligned}$$

Now, because the state $|j = 3/2, m = 3/2\rangle$ is orthogonal to the state above,

$$|j = 3/2, m = 3/2\rangle = \sqrt{\frac{3}{5}}|m_1 = 3/2, m_2 = 0\rangle - \sqrt{\frac{2}{5}}|m_1 = 1/2, m_2 = 1\rangle. \quad (4.99)$$

The Clebsch-Gordan coefficient is then

$$\langle j = 3/2, m = 3/2|j_1 = 3/2, j_2 = 1, m_1 = 3/2, m_2 = 0\rangle = \sqrt{\frac{3}{5}}. \quad (4.100)$$

Example 4.8: Neutron and Proton with Spin-Spin and Magnetic Interaction
A neutron and proton are each in an s wave of a nuclear potential. The two particles feel a spin-spin interaction,

$$V_{ss} = -\alpha \vec{S}_p \cdot \vec{S}_n.$$

The nucleons also are in region with external magnetic field of strength B,

$$V_b = -\mu_p \vec{B} \cdot \vec{S}_p - \mu_n \vec{B} \cdot \vec{S}_n.$$

In terms of α, μ_p, μ_n and B, find the four energy eigenvalues.
Solution:
This problem is made difficult by the fact that V_b is diagonal in the m_p, m_n basis, while V_{ss} is diagonal in the j, m basis. (We will omit the $s_p = 1/2, s_n = 1/2$ labels in the bras and kets to save space.)

$$\langle j, m|V_{ss}|j', m'\rangle = \delta_{j,j'}\delta_{m,m'}\frac{\alpha\hbar^2}{2}\left(j(j+1) - s_p(s_p+1) - s_n(s_n+1)\right)$$
$$\langle m_p, m_n|V_b|m_p', m_n'\rangle = -\delta_{m_p,m_p'}\delta_{m_n,m_n'}B\hbar\left(\mu_p m_p + \mu_n m_n\right)$$

To proceed further, one must choose a basis. We choose the j, m basis with the following eigenvalues,

$$|j = 1, m = 1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |j = 1, m = -1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$
$$|j = 1, m = 0\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad |j = 0, m = 0\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

In this basis V_{ss} is diagonal,

$$V_{ss} = \begin{pmatrix} -\alpha\hbar^2/4 & 0 & 0 & 0 \\ 0 & -\alpha\hbar^2/4 & 0 & 0 \\ 0 & 0 & -\alpha\hbar^2/4 & 0 \\ 0 & 0 & 0 & 3\alpha\hbar^2/4 \end{pmatrix}$$

while writing V_b requires first rewriting each of the states in the m_n, m_p basis.

$$
\begin{aligned}
|j=1, m=1\rangle &= |m_p = 1/2, m_n = 1/2\rangle, \\
|j=1, m=-1\rangle &= |m_p = -1/2, m_n = -1/2\rangle, \\
|j=1, m=0\rangle &= \frac{1}{\sqrt{2}}\left(|m_p = 1/2, m_n = -1/2 + |m_p = -1/2, m_n = 1/2\rangle\right), \\
|j=0, m=0\rangle &= \frac{1}{\sqrt{2}}\left(|m_p = 1/2, m_n = -1/2\rangle - |m_p = -1/2, m_n = 1/2\rangle\right) .
\end{aligned}
$$

If the problem involved higher angular momentum, one would have to go through the procedure of the previous exercise, utilizing raising and lowering operators, to write the j, m states in terms of the m_1, m_2 basis.
From the above form, we can see that V_b will mix the two states with with $m = 0$ as they are not eigenstates in the m_p, m_n basis.

$$
V_b = -\frac{\hbar B}{2}\begin{pmatrix} (\mu_p + \mu_n) & 0 & 0 & 0 \\ 0 & -(\mu_p + \mu_n) & 0 & 0 \\ 0 & 0 & 0 & (\mu_p - \mu_n) \\ 0 & 0 & (\mu_p - \mu_n) & 0 \end{pmatrix}.
$$

Thus, the first two eigenvalues are simple to find

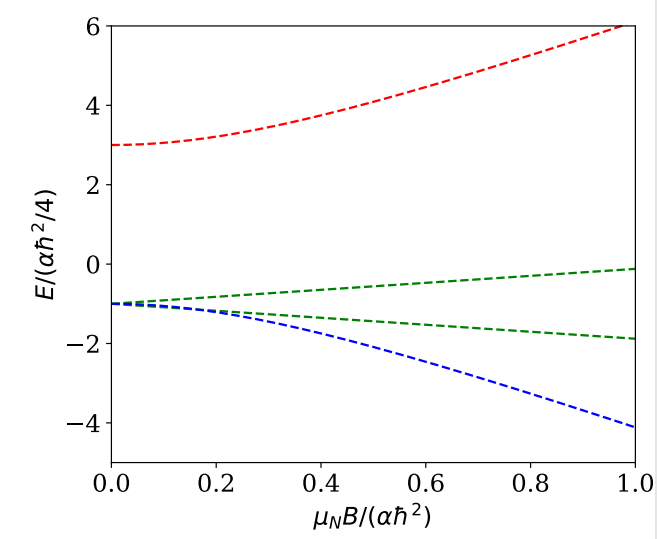
$$
\begin{aligned}
\epsilon_1 &= -\alpha\hbar^2/4 + (\mu_p + \mu_n)\hbar B/2 \\
\epsilon_2 &= -\alpha\hbar^2/4 - (\mu_p + \mu_n)\hbar B/2,
\end{aligned}
$$

while the last two eigenvalues are most easily found by writing the lower-right 2×2 submatrix in terms of σ matrices,

$$
V_{2\times 2} = \alpha\hbar^2/4 - (\alpha\hbar^2/2)\sigma_z + [(\mu_p - \mu_n)\hbar B/2]\,\sigma_x
$$

The eigenvalues of the submatrix are

$$
\epsilon_\pm = \alpha\hbar^2/4 \pm \sqrt{\alpha^2\hbar^4/4 + (\mu_p - \mu_n)^2\hbar^2 B^2/4}
$$



Energy levels of a proton and neutron experiencing a spin-spin interaction while in a magnetic field. The lower levels correspond to those of the deuteron, which has $S = 1$, whereas the upper line corresponds to an excited (unbound resonance) state.

4.15 Exercises

1. (a) Show that $\vec{r}^2 = x^2 + y^2 + z^2$ commutes with L_z.

   (b) Show that $\vec{r} \cdot \vec{p}$ commutes with L_z.

2. Any two rotations, $\vec{\alpha}$ and $\vec{\beta}$, can be written as a single rotation by $\vec{\gamma}$, which in the spin $1/2$ basis means
$$e^{i\vec{\beta}\cdot\vec{\sigma}/2} e^{i\vec{\alpha}\cdot\vec{\sigma}/2} = e^{i\vec{\gamma}\cdot\vec{\sigma}/2}.$$
Show that the equivalent angle $\vec{\gamma}$ may be written in terms of $\vec{\alpha}$ and $\vec{\beta}$ as
$$\cos(\gamma/2) = \cos(\beta/2)\cos(\alpha/2) - \hat{\beta}\cdot\hat{\alpha}\sin(\alpha/2)\sin(\beta/2)$$
$$\hat{\gamma}\sin(\gamma/2) = \cos(\beta/2)\sin(\alpha/2)\hat{\alpha} + \cos(\alpha/2)\sin(\beta/2)\hat{\beta} + \sin(\beta/2)\sin(\alpha/2)\hat{\alpha}\times\hat{\beta},$$
where $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$ are the corresponding unit vectors. Note that these relations would hold for any rotation, not just the spin $1/2$ system. Thus, they describe the rotation group.
Hints: Use the fact that $e^{i\vec{a}\cdot\sigma} = \cos(a) + i\frac{\vec{\sigma}\cdot\vec{a}}{|\vec{a}|}\sin(a)$. Also use the identity $\sigma_i\sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k$.

3. Consider the matrices,
$$S_x = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad S_y = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad S_z = \hbar\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$
These represent the rotation matrices for angular momentum $S = 1$, $S(S+1) = 2$. Note that the eigenvalues of S_z are -1,0,1 as expected.

   (a) Explicitly multiply the matrices to show that
$$[S_i, S_j] = i\hbar\epsilon_{ijk}S_k.$$
   For efficiency, just pick one of the three combinations to check.

   (b) Explicitly multiply the matrices to show that
$$\sum_i S_i^2 = 2\hbar^2\mathbb{I} = \hbar^2 S(S+1)\mathbb{I}.$$

4. Using the definition, $L_z = -i\hbar(x\partial_y - y\partial_x) = -i\hbar\partial_\phi$, express $X(\alpha) = e^{iL_z\alpha/\hbar}xe^{-iL_z\alpha/\hbar}$ in terms of x, y and z. Hint: Note that
$$e^{\alpha\partial_\phi} f(\phi) e^{-\alpha\partial_\phi} = f(\phi + \alpha).$$
because $e^{\alpha\partial_\phi}$ generates a Taylor expansion.

5. Consider the six group elements for the symmetry of the equilateral triangle listed in Sec. 4.3. As a six-by-six matrix, find the coefficients a_{ij}.

6. Using the commutation relations for angular momentum, $[L_x, L_y] = i\hbar L_z$, and the definition $L_\pm = L_x \pm iL_y$, show that
$$|\vec{L}|^2 = L_z^2 + L_+L_- - \hbar L_z.$$

7. In terms of ℓ, m_1 and m_2 find expressions for:

   (a) $\langle \ell m_1 | L_x^2 | \ell m_2 \rangle$ (Warning: this is messy)
   (b) $\langle \ell m_1 | L_x^2 + L_y^2 | \ell m_2 \rangle$

8. Consider a particle of mass m in a spherical well of radius R, where the potential is $+\infty$ for $r > R$ and zero for $r < R$.

   (a) Find the ground state energy.
   (b) Describe how one would find the energy of the first excited state of the same well.
   (c) If the particle is an electron and the radius of the well is 0.15 nm, give a numerical value for the energy of the ground state in eV.

9. (a) Estimate the ground state binding energies of the following atoms. You can use the fact that the binding energy for hydrogen is 13.6 eV, the mass of an electron is 0.511 MeV the mass of a muon is 105.7 MeV, the mass of a proton is 938.3 MeV and the charge of Pb is 82. Then scale the hydrogen values to get the desired results. The Bohr radius of H is 0.053 nm.

      i. e, Pb
      ii. μ^-, p
      iii. $e^+ e^-$
      iv. $\bar{p}, Pb$

      The mass of a muon is 205 times larger than that of an electron.

   (b) For the same cases above, find the associated Bohr radii. (Treat the Pb nucleus as a point particle)

10. For the Hydrogen atom, calculate the expectation of the operator X between the ground state and each of the four $n = 2$ states. You can express answer in terms of a_0.

11. Prove the following recurrence relation for spherical Bessel functions:

$$j_{\ell+1}(z) = -j'_\ell(z) + \frac{\ell}{z} j_\ell(z).$$

    To accomplish this, assume the equation is true and that $j_\ell(z)$ is a solution to:

$$-j''_\ell(z) - \frac{2}{z} j'_\ell(z) + \frac{\ell(\ell+1)}{z^2} j_\ell(z) = j_\ell(z).$$

    Then show that using the assumed expression for $j_{\ell+1}(z)$ will be a solution to:

$$-j''_{\ell+1}(z) - \frac{2}{z} j'_{\ell+1}(z) + \frac{(\ell+1)(\ell+2)}{z^2} j_{\ell+1}(z) = j_{\ell+1}(z).$$

    This last expression is the same differential equation as the one just above, but with $\ell \to \ell + 1$.

12. Find the Clebsch-Gordan coefficient

$$\langle \ell = 1, s = 1, j = 0, m = 0 | \ell = 1, s = 1, m_\ell = 1, m_s = -1 \rangle$$

13. Calculate the Clebsch-Gordan Coefficients $\langle \ell = 12, s = 1, j = 12, m_j = 12 | \ell = 12, s = 1, m_\ell, m_s \rangle$ for all m_ℓ and m_s.

14. An electron is in an $\ell = 1$ state of a hydrogen atom. It experiences a spin orbit interaction,

$$V_{\text{s.o.}} = \alpha \vec{L} \cdot \vec{S}$$

and also feels an external magnetic field

$$V_{\text{B}} = -\mu \vec{B} \cdot \left(\vec{L} + 2\vec{S} \right).$$

a) Using the basis

$$|J = 3/2, M = 3/2\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |J = 3/2, M = -3/2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$|J = 3/2, M = 1/2\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |J = 1/2, M = 1/2\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

$$|J = 3/2, M = -1/2\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad |J = 1/2, M = -1/2\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

write the Hamiltonian components $V_{\text{s.o.}}$ and V_{B} as 6×6 matrices. To assist you, the $|J, M\rangle$ states can be written in the $|m_\ell, m_s\rangle$ basis as

$$\begin{aligned}
|J = 3/2, m_j = 3/2\rangle &= |m_\ell = 1, m_s = 1/2\rangle, \qquad (4.101)\\
|J = 3/2, m_j = -3/2\rangle &= |m_\ell = -1, m_s = -1/2\rangle, \\
\sqrt{15/4 - 9/4 + 3/2}|J = 3/2, m_j = 1/2\rangle &= \sqrt{2}|0, 1/2\rangle + \sqrt{3/4 - 1/4 + 1/2}|-1, -1/2\rangle, \\
|J = 3/2, m_j = 1/2\rangle &= \sqrt{\frac{2}{3}}|0, 1/2\rangle + \frac{1}{\sqrt{3}}|1, -1/2\rangle, \\
|J = 1/2, m_j = 1/2\rangle &= \frac{1}{\sqrt{3}}|0, 1/2\rangle - \sqrt{\frac{2}{3}}|1, -1/2\rangle, \\
|J = 3/2, m_j = -1/2\rangle &= \sqrt{\frac{2}{3}}|0, -1/2\rangle + \frac{1}{\sqrt{3}}|-1, 1/2\rangle, \\
|J = 1/2, m_j = -1/2\rangle &= \frac{1}{\sqrt{3}}|0, -1/2\rangle - \sqrt{\frac{2}{3}}|-1, 1/2\rangle.
\end{aligned}$$

b) What are the six eigenvalues of $\boldsymbol{H}$?

15. A spin $1/2$ particle is bound to a fixed center by a spherically symmetric potential. The particle is in an $\ell = 0$ state with spin-up, i.e.

$$\Psi(\vec{r}, m) = \psi(r) \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

In terms of $\psi(r)$ and $\vec{r}$, write the matrix element for

$$\langle \vec{r}, m_s | \vec{\sigma} \cdot \vec{r} | \Psi \rangle$$

(a) for $m_s = 1/2$
(b) for $m_s = -1/2$

5 Symmetries and Conservation Laws

5.1 Continuous and Discrete Symmetries

Symmetries can be classified as either continuous or as discrete. Examples of continuous symmetries are rotational and translation symmetries while parity and time reversal are examples of discrete symmetries. Rotational and translational symmetries can be continuous because of the continuum of rotations under which the system can be rotated. However, sometimes symmetries involve rotations or translations by a finite angle. For instance, a lattice is invariant under translations of size na where n is an integer and a is the lattice spacing. A circular chain of n identically spaced identical objects is unchanged after rotation by an angle $2\pi/n$.

In classical mechanics and in field theory, one studies Noether's theorem which associates a conserved charge with any continuous symmetry. There exists an analogous relation in quantum mechanics which is perhaps easier to see. A continuous symmetry can be expressed

$$U^\dagger(\alpha)HU(\alpha) = H, \tag{5.1}$$

where H is the Hamiltonian and U represents the unitary transformation over some coordinate θ by a continuous variable α. One can write such a transformation as

$$U = e^{-iG\alpha/\hbar}, \tag{5.2}$$

where G is an operator which generates the transformation. If the operator G behaves as

$$G = i\hbar\partial_\theta, \tag{5.3}$$

then one can quickly see that

$$U = 1 + \alpha\partial_\theta + \frac{1}{2}\alpha^2\partial_\theta^2 + \cdots + \frac{\alpha^n}{n!}\partial_\theta^2 + \cdots, \tag{5.4}$$

is simply the Taylor expansion, i.e. when it operates on $f(\theta)$ it gives $f(\theta + \alpha)$.

For small α,

$$U = 1 - iG\alpha/\hbar, \tag{5.5}$$

and

$$U^\dagger(\alpha)HU(\alpha) = H - i[H, G]\alpha/\hbar, \tag{5.6}$$

which shows that H must commute with G if the Hamiltonian is to be left invariant under the transformation. Also, such a commutation implies that $\langle G\rangle$ is a constant of the motion. In the Heisenberg representation

$$\frac{d}{dt}\langle G\rangle = i\langle[H, G]\rangle/\hbar = 0 \tag{5.7}$$

Thus, the invariance of the Hamiltonian with respect to a transformation by θ is equivalent to stating that the generator of the transformation, G, is constant in time.

The degeneracies of levels is also intimately related to symmetry. For instance, if one considers an electron orbiting a proton, even with spin-orbit splitting, the Hamiltonian commutes with J^2, and because one may rotate such states into one another by changing the coordinate system, the $(2j + 1)$ states must be degenerate. If one explicitly breaks the symmetry by adding an interaction with an external magnetic field, the degeneracy is no longer expected to exist.

5.2 Parity

The parity operator $\mathbf{\Pi}$ reflects the system about some point, $x \to -x$, $y \to -y$, $z \to -z$, or equivalently three planar reflections. Parity is an example of a discrete symmetry. One can also have a system invariant under only a subset of the parity operator, e.g. a reflection about the $x = 0$ plane, $\mathbf{\Pi}_x$. Both the coordinate and momentum operators should flip sign under parity. Because $\mathbf{\Pi^2}$ should return one to the original state, eigenstates of $\mathbf{\Pi}$ should have eigenvalues of $\pm\mathbf{1}$.

The angular momentum operator,

$$\mathbf{L} \equiv \mathbf{r} \times \mathbf{p}, \tag{5.8}$$

is invariant under parity because both $\vec{r}$ and $\vec{p}$ change signs. Pseudo-vectors, which are defined by taking the cross products of two vectors, rotate like vectors but have the opposite behavior under parity. Examples of pseudo-vectors are angular momenta and the magnetic field $\vec{B}$. In fact, defining the direction of pseudo-vectors requires arbitrarily choosing a right-hand vs left-hand rule, and if parity is conserved, there should not be any behavior that differs by this choice. For instance, if some atom had its angular momentum (a pseudo-vector) oriented in the z direction by a magnetic field (also a pseudo-vector) in the z direction, the emission of a charged particle from radioactive decays should be the same parallel or anti-parallel to the applied field. After all, whether a direction is anti-parallel or parallel to the field depends on one's choice of right- vs left-handed convention, whereas the direction of a charged particle certainly defines a direction in a non-arbitrary way. In fact the electric current density is a vector. Shockingly however, experiments proposed by T.D. Lee and performed by Madame Chien-Shiung Wu Wu in 1956 showed that the weak interaction maximally violates parity by observing the decay of polarized ^{60}Co, https://en.wikipedia.org/wiki/Wu_experiment. The angular distribution of decaying electrons was strongest along the direction of the applied field, with zero probability density for a decaying electron being observed in the perfectly anti-parallel direction.

Most Hamiltonians are even under parity. In that case, the eigenstates can be either even or odd because they must be eigenstates of the parity operator. The eigenvalues are $\pm\mathbf{1}$ because $\mathbf{\Pi^2}$ must return one to the original state. If the Hamiltonian has any odd-parity terms, then the eigenstates can have mixtures of even- and odd-parity components. Further, if some original state has parity $+\mathbf{1}$, it would then develop some admixture of parity $-\mathbf{1}$ over time, which would then not be eigenstates of $\mathbf{\Pi}$. For instance, the weak interaction of an electron with a nucleus mixes the $\mathbf{2}s$ and $\mathbf{2}p$ states in the Hydrogen atom. Parity conservation refers to the state maintaining its parity, and if the Hamiltonian is not perfectly even in parity, parity is not conserved. In practice, the weak interaction violates parity, and all states have some small, often immeasurably small, admixtures of opposite parity components.

Finally, we point out that parity allows one to recognize many matrix elements as being zero. For instance, if one considers the matrix element,

$$\langle \phi | \boldsymbol{A} | \psi \rangle, \tag{5.9}$$

where the parities of the two states and the operator are $(-\mathbf{1})^{p_\phi}$, $(-\mathbf{1})^{p_A}$ and $(-\mathbf{1})^{p_\psi}$, the matrix element will be zero if $p_\psi + p_\phi + p_A$ is an odd number. This can be understood by considering what should happen when integrating over even or odd functions of the coordinate. One can

consider a matrix element for an operator $\mathcal{O}$ as being written in the form

$$\int d^3r\ \psi^*_{\alpha,\ell,m}(\vec{r})\mathcal{O}(\vec{r})\psi_{\beta,\ell',m'}. \tag{5.10}$$

For every point $\vec{r}$ there is point $-\vec{r}$. If the product $\psi^*_\alpha \mathcal{O}\psi_\beta$ is odd overall, the integral must vanish. Even if the operator involves derivative, the argument still holds. The operator $P_x = -i\partial_x$ changes the parity by a factor of -1, just the same as the operator x.

The parity of wave functions in spherical coordinates depends on the orbital quantum number,

$$\Pi|\ell, m_\ell, s, m_s\rangle = (-1)^\ell|\ell, m_\ell, s, m_s\rangle. \tag{5.11}$$

One can see this by considering the form for the $Y_{\ell m}$s. Under parity, $\cos\theta \to -\cos\theta$ whereas $\sin\theta$ does not change. The azimuthal angle $\phi \to \phi + \pi$, or equivalently $e^{im\phi} \to (-1)^m e^{im\phi}$. The harmonics $Y_{\ell,m=\ell}$ have no ϕ dependence and are even or odd functions of $\cos\theta$. The lowering operator, defined in Eq. (4.25),

$$L_\pm = -i\hbar e^{\pm i\phi}\left(\mp i\sin\theta\frac{\partial}{\partial\cos\theta} - \cot\theta\frac{\partial}{\partial\phi}\right), \tag{5.12}$$

is even under parity, as the term $e^{\pm i\phi}$ is odd as is the $\cos\theta$ dependence. Thus, all the states within a multiplet have the same parity. Further in the course, multi-particle wave functions will be considered. In that case, there is a quantum number for the total angular momentum, L. In that case the parity of the wave function will behave as $(-1)^L$. Because parity only acts on the spatial coordinates, x, y, z, the intrinsic spin part of the wave functions is unchanged under parity.

Example 5.1: Identifying Vanishing Matrix Elements by Parity
Which of the following matrix elements are zero due to parity? Here α and β refer to all other quantum numbers.

a) $\langle\alpha, \ell = 2, m = 1|x|\beta, \ell = 1, m = 0\rangle$

b) $\langle\alpha, \ell = 1, m = 1|(x^2 + y^2 - z^2)|\beta, \ell = 3, m = 1\rangle$

c) $\langle\alpha, \ell = 3, m = 0|xyz|\beta\ell = 1, m = 0\rangle$

d) $\langle\alpha, \ell = 3, m = 0|xyP_z|\beta\ell = 1, m = 0\rangle$

e) $\langle\alpha, \ell = 3, m = 0|xyP_z|\beta\ell = 1, m = 0\rangle$

Solution:

a) Can be non-zero. The operator is odd under parity, and the bra and the ket have different parities.

b) Can be non-zero. In this case the operator is even under parity and the bra and ket have the same parities.

c) Must be zero. The bra and kets have the same parity, but the operator is odd under parity.

d) Must be zero. The operator has the same parity as (c).

e) Can be non-zero. The operator $\boldsymbol{L_z}$ has even parity, so the operator has even parity and can link states of the same parity.

5.3 Time Reversal

Time reversal is an odd kind of symmetry. It suggests that a motion picture of a physical event could be run in reverse without the viewer being able to tell something is wrong. Of course, this does not apply whenever dissipation (friction) is included. In classical mechanics the motion of a particle in a potential $\boldsymbol{V(r)}$ should be equally valid when watched in reverse. In electrodynamics, the reverse motion should be fine as long as the direction of the magnetic field is reversed. Summarizing time reversal behavior of particles in classical electrodynamics,

$$\vec{\boldsymbol{r}} \rightarrow \vec{\boldsymbol{r}},\ \vec{\boldsymbol{v}} \rightarrow -\vec{\boldsymbol{v}},\ \mathbf{E} \rightarrow \mathbf{E},\ \mathbf{B} \rightarrow -\mathbf{B},\ \mathbf{j} \rightarrow -\mathbf{j}, \boldsymbol{\rho} \rightarrow \boldsymbol{\rho}. \tag{5.13}$$

should lead to an equally valid evolution. To understand the signs above, imagine a charged particle moving in an electric field from times $\boldsymbol{t_0}$ to $\boldsymbol{t_f}$. If at $\boldsymbol{t_f}$ you reversed the velocity, the particle would trace out the identical trajectory. However, to get a particle to retrace its trajectory in a magnetic field, one would have to flip the sign of the magnetic field. The charge density is not a function of time reversal, and because the velocity clearly changes sign, and because the current density is $\boldsymbol{\rho\vec{v}}$, clearly the current density is odd under time reversal.

The electric potential Φ and the vector potential $\vec{\boldsymbol{A}}$ appear in the combinations, $\boldsymbol{H_0 - e\Phi}$ and $\vec{\boldsymbol{P}} - \boldsymbol{e}\vec{\boldsymbol{A}}/\boldsymbol{c}$, so one expects Φ and $\vec{\boldsymbol{A}}$ to have the same time-reversal properties as the Hamiltonian and as the momentum, respectively, i.e. Φ is even and $\vec{\boldsymbol{A}}$ should be odd. In fact, for all Lorentz four-vectors, the three spatial components must have the opposite behavior under time reversal as the "zero$^{\mathbf{th}}$" component.

In quantum mechanics, we expect the following behavior under time reversal,

$$\boldsymbol{t} \rightarrow -\boldsymbol{t},\ \mathbf{r} \rightarrow \mathbf{r},\ \mathbf{p} \rightarrow -\mathbf{p}. \tag{5.14}$$

The gradient operator and $\boldsymbol{\partial_t}$ should be even and odd, respectively, under time reversal. However $\vec{\boldsymbol{P}} = -\boldsymbol{i\hbar\nabla}$ and $\boldsymbol{H} = \boldsymbol{i\hbar\partial_t}$ are odd and even respectively if they are to represent the momentum and energy. Thus, for quantum mechanics, the time reversal operator must also involve taking the complex conjugate. Under time reversal,

$$\begin{aligned} \boldsymbol{i\hbar\partial_t} &\rightarrow \boldsymbol{i\hbar\partial_t}, \\ -\boldsymbol{i\hbar\nabla} &\rightarrow \boldsymbol{i\hbar\nabla}. \end{aligned} \tag{5.15}$$

One may also see that time reversal entails taking the complex conjugate by considering Schrödinger's wave equation,

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi(x,t) + V(x)\Psi(x,t) = i\hbar\frac{\partial}{\partial t}\Psi(x,t). \tag{5.16}$$

If $\Psi(x,t) = \psi(x)e^{-iEt}$ is a solution, taking the complex conjugate and reversing the time also gives a solution to the Hamiltonian. In fact, because $\boldsymbol{H}$ is real one may write $\psi(x)$ as a purely real function or purely imaginary function. A purely real function $\psi(x)$ would be even under time reversal symmetry while a purely imaginary function would be odd. For instance a plane wave solution, $\psi(x)$ is neither even nor odd as it has both real and imaginary parts. In that case taking the complex conjugate changes the direction of the momentum, thus signifying the motion reversal mentioned before.

Considering the example of plane waves,

$$\Psi(x,t) = e^{-iEt+ikx}. \tag{5.17}$$

By both taking the complex conjugate and reversing the sign of t, one obtains a new solution where the momentum is reversed but the energy remains the same. These are eigenstates of the Hamiltonian, but are not eigenstates of the time-reversal operator, even though the time-reversal operator commutes with the Hamiltonian. This puzzle resolves itself by realizing that the time-reversed state has the same energy, so one can make combinations,

$$\Psi_{\pm}(x,t) = \frac{1}{\sqrt{2}}\left(e^{-iEt+ikx} \pm e^{-iEt+ikx}\right),$$

which are eigenstates of both the Hamiltonian, and of the time-reversal operator, with eigenvalues $\pm\mathbf{1}$.

Of course, when one gets away from the Schrödinger equation, there are many Hamiltonians which are not real but are Hermitian, e.g. those that involve $\boldsymbol{\sigma}_y$. Nonetheless, there usually exists a time-reversal symmetry, but expressing the symmetry may be more complicated than merely taking the complex conjugate. In particle physics, the symmetry of switching times is linked to the existence of solutions with opposite energy (anti-particles). In that case the operation which includes taking the complex conjugate is associated with finding negative-energy or antiparticle solutions, but that is material for another course.

The time reversal operator acts only to the right because it entails taking the complex conjugate. In terms of bras and kets the time-reversed matrix elements satisfy the relations below, using the notation, $|\tilde{\alpha}\rangle = \Theta|\alpha\rangle,\ \tilde{\alpha}_i = \alpha_i^*$.

$$\langle\tilde{\beta}|\tilde{\alpha}\rangle = (\beta_i^*)^*\alpha_i^* = \alpha_i^*\beta_i = \langle\alpha|\beta\rangle = \langle\beta|\alpha\rangle^*. \tag{5.18}$$

Similarly, the expectation of operators satisfies the relations,

$$\langle\tilde{\beta}|\Theta A\Theta^{-1}|\tilde{\alpha}\rangle = \langle\tilde{\beta}|\Theta A|\alpha\rangle = (\beta_i^*)^* A_{ij}^*\alpha_j^* = \langle\alpha|A^\dagger|\beta\rangle = \langle\beta|A|\alpha\rangle^*, \tag{5.19}$$

or is equivalent to saying that the time reversed operator $\Theta A\Theta^{-1}$ sandwiched between time-reversed states gives the complex conjugate of the same matrix element without the time reversals.

Most operators of interest are either even or odd under time reversal.

$$\Theta B\Theta^{-1} = \pm B \tag{5.20}$$

Clearly an operator that is a linear combination of an odd and an even operator would be neither. Examples of odd and even operators are:

$$\Theta\vec{P}\Theta^{-1} = -\vec{P},\ \ \Theta\vec{A}\Theta^{-1} = -\vec{A},\ \ \Theta\vec{r}\Theta^{-1} = \vec{r},\ \ \Theta\vec{L}\Theta^{-1} = -\vec{L}, \tag{5.21}$$

where $\vec{A}$ is the electromagnetic vector potential.

Note that the commutation relations,

$$[x, p] = i\hbar\,, \;\; [L_i, L_j] = i\hbar\epsilon_{ijk}L_k, \tag{5.22}$$

have apparently different behaviors under reversal on the left and right-hand sides of the equations, until one remembers to consider the $i\hbar$, which flips sign under the time-reversal operator. Notice that the raising and lowering operator for a harmonic oscillator, $x + ip$, is even under time reversal.

Hamiltonians are usually invariant under time reversal. Here we list a few terms which might appear in a Hamiltonian and discuss whether they violate time reversal or parity.

1. $\mathbf{p^2}/2m$ is invariant under both.
2. $\mathbf{p}\cdot\mathbf{r}$ is invariant under parity but not time reversal.
3. $\mathbf{L}\cdot\mathbf{p}$ is invariant under time reversal but not parity.
4. $\mathbf{S}\cdot\mathbf{B}$ and $\mathbf{p}\cdot\mathbf{A}$ are invariant under both.

5.4 Time Reversal and Angular Momentum

The time reversal operator acts in a surprisingly complex manner when operating on eigenstates of angular momentum. For integer-spin particles, one can understand the behavior by considering properties of the $Y_{\ell,m}$s. By taking the complex conjugates, one sees

$$\Theta|j, m\rangle = (-1)^m|j, -m\rangle. \tag{5.23}$$

However, we should realize that the $(-1)^m$ is basically the result of a phase convention and should not be taken too seriously.

The case of spin 1/2 particles is more surprising. In this case the spins cannot be represented by $Y_{\ell,m}$s and one must instead consider the two-component system where $|+\rangle$ and $|-\rangle$ refer to spin-up and spin-down with the direction being chosen along the z axis. In this basis an arbitrary normalized state, with spin in some arbitrary direction, may be written,

$$|\alpha\rangle = e^{i\delta_1}\begin{pmatrix}\cos\theta/2\\ e^{i\gamma_1}\sin\theta/2\end{pmatrix}. \tag{5.24}$$

Because the time-reversed state must be a state with opposite spin, it must be orthogonal. The time reversal operator must create the orthogonal state,

$$\Theta|\alpha\rangle = e^{i\delta_2}\begin{pmatrix}-e^{-i\gamma_1}\sin\theta/2\\ \cos\theta/2\end{pmatrix} \tag{5.25}$$

The phases δ_1 and δ_2 are arbitrary, and labeling their difference as $\eta \equiv \delta_2 - \delta_1$ we see that Θ must equal

$$\Theta = e^{i\eta}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}K, \tag{5.26}$$

where K is an operator which takes the complex conjugate of any quantities to the right. The surprising property of Θ is viewed by squaring the operator,

$$\Theta^2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^2 = -1, \tag{5.27}$$

independent of the choice of η.

Making an arbitrary choice of $e^{i\eta} = -i$, one sees that

$$\Theta|+\rangle = i|-\rangle \ , \ \ \Theta|-\rangle = -i|+\rangle. \tag{5.28}$$

This allows one to summarize both the integral and half-integral results with

$$\Theta|j,m\rangle = i^{2m}|j,-m\rangle \tag{5.29}$$

Thus, performing two time-reversal operations does not return the same state in the case of half-integral spin, but it does for the case of integral spin.

To make the time-reversal operator in spin 1/2 systems seem a bit more peculiar it is often written as

$$\Theta = \sigma_y K, \tag{5.30}$$

which makes the operator look odd because the y direction appears to be preferentially singled out. However, this has nothing to do with the y direction, but only with the fact that the only the operator which flips spin must change the place of the two components and change one sign to make an orthogonal state. One could have associated any of the three sigma matrices with the x, y and z directions, but the time reversal operator must be proportional to the anti-symmetric off-diagonal matrix.

Note that the eigenstates of L^2 and L_z are not eigenstates of the time-reversal operator, unless the projection m equals zero, because they have the form $e^{im\phi}$. However, one can make eigenstates of the time reversal operator by combining the states $|m\rangle$ and $|-m\rangle$, e.g. $(|m\rangle \pm |-m\rangle)/\sqrt{2}$.

5.5 A Particle in a Periodic Potential

A second class of discrete symmetries involves particles moving in a periodic lattice, which for our purposes will be considered to be a one-dimensional chain. Two- and three-dimensional lattices are more difficult to describe, but many basic features, such as the existence of band structure, occur in all dimensionalities. The translation operator, τ_a, translates the wave function by a distance a, where the lattice separation is a. Translational invariance implies that τ_a commutes with the Hamiltonian,

$$\tau_a H \tau_a^{-1} = H \ \text{ or } \ V(x+a) = V(x). \tag{5.31}$$

One thus expects the solution to be an eigenstate of the translation operator,

$$\tau_a \psi_k(x) = \psi_k(x+a) = e^{ika}\psi_k(x), \tag{5.32}$$

where the eigenvalue is e^{ika}. We label the eigenvalue by k, but emphasize that k is not the eigenvalue of the momentum operator and $|k\rangle$ is not an eigenstate of the momentum operator.

Instead, k is a label that denotes that the eigenvalue of τ_a is e^{ika}. The translation operator τ_a commutes with the Hamiltonian, but it is not Hermitian. Thus, the eigenvalue can be complex. If the physics is to be unchanged by translation, the state must change only by a phase factor. Hence, the form of the eigenvalue can be chosen as e^{ika} with ka being real. The value ka can take on any value, but must be in a range of $\Delta ka = 2\pi$, e.g. $-\pi < ka < \pi$. Any value of ka outside this range gives the same eigenvalue as one inside the range where you pick a value within the range that differs by integral numbers of 2π. For a given ka, one can have a multitude of eigenstates, as will be illustrated below.

The Kronig-Penney model is enormously useful in illustrating how band structure emerges from the periodicity of a lattice. It describes a one-dimensional geometry with delta-function potentials separated by a,

$$V(x) = \sum_{n=-\infty}^{\infty} \beta\delta(x - na). \tag{5.33}$$

For positions $0 < x < a$ there is no potential energy, so if the eigenenergy is $\hbar^2q^2/2m$, the form of the wave function is

$$\psi(x) = e^{iqx} + Be^{-iqx}, \tag{5.34}$$

with boundary conditions,

$$\psi(a) = e^{ika}\psi(0)\ , \quad \frac{d}{dx}\psi(x)|_{x=a-\epsilon} - e^{ika}\frac{d}{dx}\psi(x)|_{x=\epsilon} + \frac{2m\beta}{\hbar^2}\psi(a) = 0. \tag{5.35}$$

The boundary conditions become

$$\begin{aligned} e^{iqa} + Be^{-iqa} &= e^{ika}(1 + B) \\ iq\left(e^{iqa} - Be^{-iqa} - e^{ika} + e^{ika}B\right) &= -\frac{2m\beta}{\hbar^2}e^{ika}(1 + B) \end{aligned} \tag{5.36}$$

Eliminating B, one can find a transcendental expression for q.

$$p\sin(qa) + 2q\cos(qa) - 2q\cos(ka) = 0, \tag{5.37}$$

where $p \equiv 2m\beta/\hbar^2$. Because the solution only depends on $\cos(ka)$, there are solutions for $-\pi < ka < \pi$ with the solutions symmetric about $k = 0$. However, for any k there are a variety of solutions q to the transcendental equation. Plotting the solutions as a function of k yields bands. In the limit $p \to 0$, the solutions are given by $q = k$, otherwise they are shifted, especially near the points where $\cos(ka) = \pm 1$ as seen in the example in Fig. 5.1.

The bands refer to the energies that give solutions. Those energies, $\hbar^2q^2/2m$, can be found for the values of qa, which are plotted along the vertical axis in Fig. 5.1. There are values of qa, or equivalently of the energy, that are not solutions for some value of ka. These are called gaps, and the gap typically refers to the jump in energy between bands. If the strength of the potential, β were set to zero, then the term proportional to p in Eq. (5.37) would vanish, and one would find

$$\begin{aligned} q(\beta = 0) &= k \pm \frac{2n\pi}{a}, \\ E(k) &= \frac{\hbar^2q^2}{2m} = \frac{\hbar^2(k \pm 2n\pi/a)^2}{2m}, \end{aligned} \tag{5.38}$$

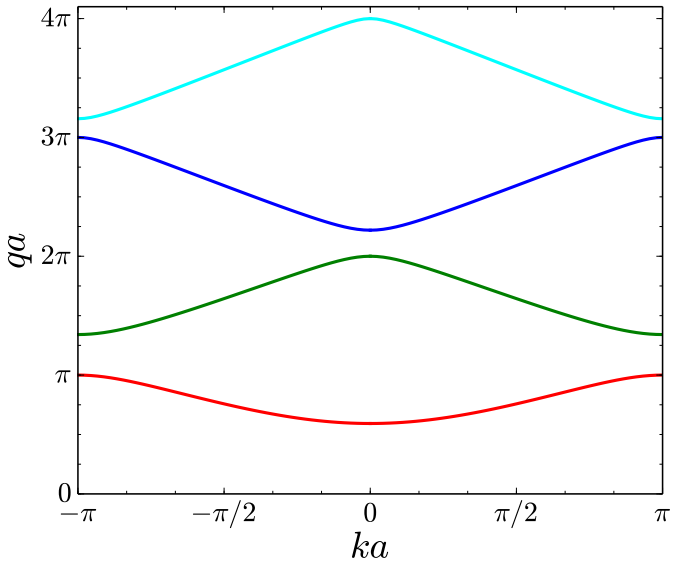


Figure 5.1: First four bands of the Kronig Penny model from Eq. (5.37) where the lattice spacing is a, and for a coupling where $pa = 5$. Energies would be given by $\hbar^2 q^2/2m$.

where $n = 0$ refers to the first band. So, by mapping $k \to k \pm 2n\pi/a$, the energies would be the same as for free particles.

The structure of bands plays a critical role in understanding the electronic properties of solids. Depending on the density of electrons, the bands are filled up to some level. If that level occurs in the middle of a band, it is easy for electrons to become excited within the band and arrange themselves with more positive momenta than negative momenta and yield a current. Such materials can be good conductors. If the density of electrons is such that the cutoff for filling levels is between bands, then it requires energy to overcome the band gap to create currents. These materials are insulators. If the cutoff is between two bands with a small gap (compared to the temperature), the material is a semi-conductor. This concept extends to multiple dimensions, and the band structure will depend on the geometrical symmetries of the underlying ionic lattice.

5.6 Object with Discrete Rotational Symmetry

Another example of a discrete symmetry is that of the rotations of a symmetrical object. Unlike the Kronig-Penney model, there is no particle moving in a symmetric potential. Rather, the symmetry is the object itself. Even an object with no noticeable symmetry will be invariant under rotations of 2π. In two-dimensions, an oval is invariant under rotations by π, and a square is invariant under rotations of $\pi/2$. The wave function of such an object would use and angle ϕ to describe the position, but the angle would only have values between zero and π for the oval or between zero and $\pi/4$ for the square. Rotations for larger angles would be equivalent to rotations within the smaller range. The wave functions would still behave as $e^{im\phi}$, but with the constraint that

$$e^{im\Phi} = 1 \tag{5.39}$$

when Φ corresponds to a rotation that brings one back to the original state. For the oval, where $\Phi = \pi$, this then constrains m to the values $0, 2, 4, 6 \cdots$, whereas for the square $m =$

$\mathbf{0, 4, 8, 12, 16 \cdots}$. This requires a perfect symmetry. For example, if the square had distinguishable particles located on each corner one could distinguish between a rotation of $\boldsymbol{\pi/2}$ and rotation of zero. Thus, the values of $\boldsymbol{m}$ would return to $\mathbf{0, 1, 2, 3 \cdots}$. Even if the particles on each had the exact same mass and charge, the values would be $\mathbf{0, 1, 2, 3}$. Thus, the first rotational state is at a higher energy if the symmetry is higher. In fact if the object was perfectly circular, $\boldsymbol{n \to \infty}$, one could not rotate it.

This has implications for nuclear and atomic physics. Many nuclei are elliptical and have a reflection symmetry, $\psi(\hat{\boldsymbol{n}}) = \psi(-\hat{\boldsymbol{n}})$. These are therefore invariant under rotations by $\boldsymbol{\pi}$, and only even values of $\boldsymbol{n}$ are allowed. For three-dimensional objects only even $\boldsymbol{\ell}$ are allowed for elliptic objects, and the sequence of energy levels is constructed with $\boldsymbol{\ell = 0, 2, 4 \cdots}$. These are referred to as rotational bands, with the term "band" having a different meaning than in the Kronig-Penney model. To emphasize the importance of symmetry, one could consider an $\boldsymbol{O_2}$ molecule. If one of the atoms were a different isotope, then both even and odd values of $\boldsymbol{\ell}$ would be permitted, but if both were the same isotope only even values of $\boldsymbol{\ell}$ would be permitted.

Example 5.2: Rotating Chain
Here, we consider an example of a circular chain of mass $\boldsymbol{M}$ and radius $\boldsymbol{R}$ with $\boldsymbol{n}$ equally spaced identical particles fixed to the chain that can rotate together. The chain exists in two dimensions. Find the rotational energies.
Solution:
The chain is invariant under rotations of $\boldsymbol{2\pi/n}$. Because this is rotation in a plane, the position of the chain is specified by a single angle

$$\boldsymbol{0 < \phi < 2\pi/n.}$$

Any rotation outside this range would be equivalent to a rotation inside the range. The part of the Hamiltonian representing the rotational kinetic energy would be of the form

$$\boldsymbol{H_K = \frac{L^2}{2MR^2} = -\frac{\hbar^2}{2MR^2}\partial_\phi^2.}$$

Because the object rotated by $\boldsymbol{2\pi/n}$ is indistinguishable, the boundary condition must be that the wave function is returned to the same state, within a phase, by a rotation by $\boldsymbol{2\pi/n}$,

$$\boldsymbol{\psi(\phi + 2\pi/n) = \psi(\phi).}$$

This differs from the Kronig-Penny model in that the object carries the symmetry, where in the Kronig-Penny model the object was a point particle moving in a periodic potential. When this chain is rotated by $\boldsymbol{2\pi/n}$ one has the same state, but when an electron is translated a distance $\boldsymbol{a}$ in a periodic lattice, it is indeed a different state.

$$\boldsymbol{\psi_m(\phi) = e^{im\phi}\,,\ \ m = \cdots, -2n, -n, 0, n, 2n, \cdots,}$$
$$\boldsymbol{E_m = \frac{\hbar^2 m^2}{2MR^2}.}$$

If one link in the chain were different, none of these arguments would apply as only rotations by $\boldsymbol{2\pi}$ would return one to the original state and $\boldsymbol{m}$ could be any integer.

5.7 Exercises

1. Let $\mathcal{T}_{\vec{d}}$ denote the translation operator (displacement vector $\vec{d}$); $\mathcal{D}(\hat{n},\phi)$, the rotation operator; and π the parity operator. Which, if any, of the following pairs commute? Why?

   (a) $\mathcal{T}_{\vec{d}}$ and $\mathcal{T}_{\vec{d}'}$ ($\vec{d}$ and $\vec{d}'$ are in different directions.)
   (b) $\mathcal{D}(\hat{n},\phi)$ and $\mathcal{D}(\hat{n}',\phi')$ ($\hat{n}$ and $\hat{n}'$ are in different directions.)
   (c) $\mathcal{T}_{\vec{d}}$ and Π.
   (d) $\mathcal{D}(\hat{n},\phi)$ and Π.

2. Because of weak (neutral-current) interactions there is a parity-violating potential between the atomic electron and the nucleus as follows:

$$\boldsymbol{V} = \boldsymbol{\lambda}\left[\boldsymbol{\delta}^3(\vec{r})\mathbf{S}\cdot\mathbf{p} + \mathbf{S}\cdot\mathbf{p}\boldsymbol{\delta}^3(\vec{r})\right]$$

   where $\mathbf{S}$ and $\mathbf{p}$ are the spin and momentum operators of the electron, and the nucleus is assumed to be situated at the origin. As a result, the ground state of an alkali atom, usually characterized by $|\boldsymbol{n},\boldsymbol{\ell},\boldsymbol{j},\boldsymbol{m}\rangle$ actually contains tiny contributions from other eigenstates as follows

$$|\boldsymbol{n},\boldsymbol{\ell},\boldsymbol{j},\boldsymbol{m}\rangle \to |\boldsymbol{n},\boldsymbol{\ell},\boldsymbol{j},\boldsymbol{m}\rangle + \sum_{\boldsymbol{n}',\boldsymbol{\ell}',\boldsymbol{j}',\boldsymbol{m}'} \boldsymbol{C}_{\boldsymbol{n}',\boldsymbol{\ell}',\boldsymbol{j}',\boldsymbol{m}'}|\boldsymbol{n}',\boldsymbol{\ell}',\boldsymbol{j}',\boldsymbol{m}'\rangle$$

   On the basis of symmetry considerations *alone*, what can you say about $(\boldsymbol{n}',\boldsymbol{\ell}',\boldsymbol{j}',\boldsymbol{m}')$ which give rise to non-vanishing contributions?

3. Suppose a spinless particle is bound to a fixed center by a potential $\boldsymbol{V}(\vec{r})$ so asymmetrical that no two energy levels are degenerate. Using time-reversal invariance prove

$$\langle\mathbf{L}\rangle = \mathbf{0}.$$

   for any energy eigenstate. Use the fact each eigenwave function $\psi(\vec{r})$ must be an eigenstate of the time reversal operator with eigenvalue $e^{i\gamma}$, thus $\psi^*(\vec{r}) = e^{i\gamma}\psi(\vec{r})$. Also use the fact that $\langle\boldsymbol{\alpha}|\boldsymbol{L}_i|\boldsymbol{\alpha}\rangle$ is real because $\boldsymbol{L}_i$ is a Hermitian operator. (This is known as **quenching** of orbital angular momentum.)

4. Consider the time-reversal operator for spin-1/2 particles, $\boldsymbol{\Theta} = \boldsymbol{\sigma}_y\boldsymbol{K}$, where $\boldsymbol{K}$ takes the complex conjugate of all quantities to its right. Show that $\boldsymbol{\Theta}$ commutes with the rotation operator,

$$\boldsymbol{\mathcal{R}}(\vec{\boldsymbol{\theta}}) = \cos(\boldsymbol{\theta}) + \boldsymbol{i}\vec{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{\theta}}\sin(\boldsymbol{\theta}).$$

5. Consider a particle of mass $\boldsymbol{M}$ confined to a two-dimensional circle of radius $\boldsymbol{R}$.

   (a) Write down the Schrödinger equation for the wave function $\psi(\phi)$, where the potential depends only on ϕ, and radial motion is ignored.
   (b) Assuming the potential is periodic,

$$\boldsymbol{V}(\boldsymbol{\phi} + \boldsymbol{2\pi/N}) = \boldsymbol{V}(\boldsymbol{\phi}),$$

   where $\boldsymbol{N}$ is an integer. Write the boundary condition relating $\psi(\phi)$ and $\psi(\phi+2\pi/N)$, where the eigenvalue of the rotation operator, $\boldsymbol{\mathcal{R}}(\boldsymbol{2\pi/N})$, is $\boldsymbol{e}^{i\gamma}$. What values of γ are allowed?

(c) Assume the potential ,

$$V(\phi) = \beta \sum_{k=1,N} \delta(\phi - 2\pi k/N),$$

Assume the wave function has the form,

$$\psi(\phi) = e^{im\phi} + Be^{-im\phi}, \;\; 0 < \phi < 2\pi/N,$$

where m is not necessarily an integer. Find a transcendental expression for m in terms of β, M, γ and n. Hint: Note the similarity to the Kronig-Penny model, where the solution in Eq. (5.37) translates to this problem with $qa \to m\alpha$, $ka \to \gamma$, and $a \to \alpha$, with $\alpha = 2\pi/N$.

6 Approximation Methods

Physics is rarely solved exactly, and the art of approximating the correct answer is the essence of any physics problem. Even if a problem cannot be solved exactly, one may be able to satisfactorily approximate the true solution. A wide range of approximation methods exist, that vary widely in the circumstances under which they become accurate. For instance, the WKB approximation provides a crude estimate of tunneling probabilities that is valid when the potential changes slowly with position. Variational methods encompass a wide variety of approaches and becomes exact when the true wave function can be well described with some form involving a few adjustable parameters. For transitions, it is good to understand whether the Hamiltonian is changing quickly or slowly, as if the time scale describing the change is either much shorter or much longer than any characteristic frequencies of the problem, one can apply either the sudden or adiabatic approximation. Finally, if a Hamiltonian differs by a small amount from one which is exactly solvable, the exact solution can be expressed as a converging sum of perturbative corrections. This chapter covers five different classes of approximation:

1. The WKB Approximation
2. Variational Theory
3. The Sudden Approximation
4. Stationary-State Perturbation Theory
5. Time-Dependent Perturbation Theory

The choice of approximation method depends both on the justifiability of the approximation and on what is being asked. For instance, time-dependent perturbation theory is applied for problems where a transition is being studied, such as a decay or a scattering. For addressing questions about a stationary states, e.g. how is the binding energy of the ground state affected by a small additional potential, one might apply stationary-state perturbation theory or a variational method. The methods presented here by no means constitute an exhaustive list. Lattice gauge theory the method becomes exact in the limit that the space time mesh becomes infinitely fine.

6.1 The WKB Approximation

The WKB (Wentzel-Kramer-Brillouin) approximation is a useful method for estimating wave functions and tunneling probabilities for smooth potentials or for potentials with only a few discontinuities. The WKB approximation for a wave function can be written as:

$$\begin{aligned} \psi(x) &= A_+(x)e^{i\phi(x)} + A_-(x)e^{-i\phi(x)} \\ \phi(x) &= \int^x dx' p(x')/\hbar, \end{aligned} \tag{6.1}$$

where the lower limit of the integral is absorbed by the arbitrary phases in A_+ and A_-. The function $p(x)$ is defined by

$$p(x) \equiv \sqrt{2m(E - V(x))}, \tag{6.2}$$

and can be thought of the momentum of a classical particle with energy E at position x.

To assess the accuracy of the approximation one can apply the Schrödinger equation to the assumed form for ψ. Here we use only the A_+ term.

$$\begin{aligned}(H-E)\psi(x) &= -\hbar^2\frac{\partial^2}{\partial x^2}\psi(x) - (E-V)\psi(x) \\ &= \left\{\frac{p(x)^2}{2m} - E - V(x)\right\} A_+(x)e^{i\phi(x)} \\ &+ \frac{i\hbar}{2m}\left\{2p(x)\frac{\partial}{\partial x}A_+(x) + A_+(x)\frac{\partial}{\partial x}p(x)\right\}e^{i\phi(x)} \\ &- \frac{\hbar^2}{2m}\left\{\frac{\partial^2}{\partial x^2}A_+(x)\right\}e^{i\phi(x)}.\end{aligned} \tag{6.3}$$

The first term disappears from our choice of $p(x)$ while the second term will vanish if we choose

$$A_+(x) \propto p(x)^{-1/2}. \tag{6.4}$$

The last term does not vanish, but is neglected in the limit that $\hbar$ is small as it is proportional to $\hbar^2$. One can understand the accuracy of the approximation by taking the second derivative of $A_+(x)$ in the last term and comparing it to the other terms. One then sees that the approximation becomes exact when characteristic length scales of the the potential are much longer than $\hbar/p(x)$.

Physically, one may understand the $p^{-1/2}$ dependence of A_+ by realizing that the WKB approximation has no reflection associated with it, at least when $E > V$ and $p(x)$ is real. Thus, conservation of flux requires $p|A|^2$ to be constant.

To estimate the binding energy, one sets the phase $\phi(x)$ to be zero at one turning point (a point where $p(x) = 0$), then solves for $\phi(x)$ at the other turning point. By finding an energy for which the phase changes by π one then has a solution. For one may then combine such a solution with the A_- solution to return to the original turning point incurring a net phase change of 2π.

Example 6.1: Ground State of Harmonic Oscillator in WKB approximation
Consider a particle of mass m in the one-dimensional harmonic oscillator potential,

$$V(x) = \frac{1}{2}m\omega^2x^2. \tag{6.5}$$

estimate the energies of the eigenstates.
Solution:
For bound states, one typically applies the WKB approximations between the classical turning points. This involves assuming that the WKB phase is a multiple of π between the points defined by $V(x) = E$. In this case one fourth of $2n\pi$ must result from integrating $\int k(x)dx$ from zero to the turning point.

$$\frac{n\pi}{2} = \frac{\sqrt{2m}}{\hbar}\int_0^a dx\sqrt{\frac{1}{2}m\omega^2a^2 - \frac{1}{2}m\omega^2x^2}, \tag{6.6}$$

where the turning point a depends on the unknown ground state energy, $E = m\omega^2a^2/2$.

Solving the equation above for a gives

$$\begin{aligned} a^2 &= \frac{2n\hbar}{m\omega} \\ E &= n\hbar\omega, \end{aligned} \tag{6.7}$$

which is correct aside from an extra $\hbar\omega/2$. Because it was assumed that the wave function vanished at the turning point (where $p(x) = 0$), rather than allowing the wave function to gradually decay in the region where $p(x)$ is imaginary, one expects it to consistently give an over-prediction of the energies.

Because other methods tend to be much more accurate, the WKB approach is rarely used to estimate ground state energies. The most common use of the WKB approximation is to estimate tunneling probabilities. In this case the wave functions have exponentially growing and decaying amplitudes.

$$\begin{aligned} \psi(x) &= A_+(x)e^{\phi(x)} + A_-(x)e^{-\phi(x)} \\ \phi(x) &= \int^x dx' q(x')/\hbar, \end{aligned} \tag{6.8}$$

where the approximation is applied to tunneling regions, $V(x) > E$. The function $q(x)$ is defined as

$$q(x) \equiv \sqrt{2m(V(x) - E)}, \tag{6.9}$$

In fact, one usually ignores the x dependence of the amplitude and merely states that the tunneling probability form going from turning point a to turning point b is

$$P_{a\to b} \approx \exp\left\{-\frac{2}{\hbar}\int_a^b dx\sqrt{2m(V(x) - E)}\right\}, \tag{6.10}$$

with the factor of two coming from squaring the amplitude.

Example 6.2: Tunneling Through Barrier
Consider a particle in a potential,

$$V(x) = \left\{ \begin{array}{ll} \infty, & x < 0 \\ 0, & 0 < x < L \\ V_0, & L < x < 2L \\ 0, & x > 2L \end{array} \right\}$$

where the barrier, V_0 is large. Estimate the tunneling probability of a particle stuck in the well.
Solution:

First, estimate the energy of the ground state ignoring any penetration,

$$\begin{aligned}\pi &= \int_0^L dx\, \sqrt{2mE}/\hbar \\ &= \frac{L}{\hbar}\sqrt{2mE}, \\ E &= \frac{\hbar^2\pi^2}{2mL^2}.\end{aligned}$$

Next, estimate the rate at which the particle collides with the right side to try and escape. The distance back and forth across the well is $2L$, and if the velocity is $\sqrt{2E/m}$, the rate is

$$\Gamma_{try} = \frac{v}{2L} = \frac{\hbar\pi}{2mL^2}.$$

Finally, one can apply the WKB approximation to estimate the tunneling probability,

$$\begin{aligned}P_{\text{tunnel}} &= \exp\left\{-\frac{2}{\hbar}\int_L^{2L} dx\, p(x)\right\}, \\ p(x) &= \sqrt{2m(V(x)-E)} = \sqrt{2m(V_0-E)}, \\ P_{\text{tunnel}} &= \exp\left\{-\frac{2L}{\hbar}\sqrt{2m(V_0-E)}\right\}, \\ &= \exp\left\{-\frac{2L}{\hbar}\sqrt{2m(V_0-\frac{\hbar^2\pi^2}{2mL^2})}\right\}.\end{aligned}$$

The factor of two in the exponential is from squaring the wave function, which is assumed to be unity at $x = L$. The estimate for the decay rate is then

$$\Gamma = \Gamma_{\text{try}} P_{\text{tunnel}}.$$

Every step of the WKB approximation is crude, but having an estimate of the decay rate through a barrier can be useful even if the accuracy is only within an order of magnitude given that decay rates for processes like fission can range from 10^{-20} seconds to many times the age of the universe.

6.2 Variational Theory

Variational calculations are used to calculate ground state wave functions by using the simple fact that all states must have energies greater or equal to that of the ground state. Thus, by writing a state in terms of some parameters α_i, and minimizing $E = \langle\vec{\alpha}|H|\vec{\alpha}\rangle$ with respect to the parameters α_i, one knows that the energy can never fall below the true ground state energy. The minimization procedure therefore can be used to estimate the ground state energy and ground state wave functions.

Example 6.3: Variational Estimate of Ground State Energy of Hydrogen Atom
Consider the Coulomb potential,

$$V(r) = -\frac{e^2}{r}.$$

We suppose that we were lucky and guessed an exponential form for ground-state hydrogen atom wave function by assuming a trial form,

$$\psi(r) = \frac{2}{\sqrt{4\pi a^3}} e^{-r/a},$$

where a is the variational parameter and the prefactor was chosen to normalize the wave function. One would then minimize the expectation of $\langle H \rangle$ with respect to a.

$$\langle H \rangle = 4\pi \int r^2 dr\, \psi^*(r) \left[-\frac{\hbar^2}{2m}(\partial_r^2 + (2/r)\partial_r) - \frac{e^2}{r} \right] \psi(r),$$

and after doing the integrals,

$$\langle H \rangle = \frac{\hbar^2}{2ma^2} - \frac{e^2}{a}.$$

The minimization, $(\partial/\partial a)\langle H \rangle = 0$, yields

$$a = \frac{\hbar^2}{me^2},$$
$$\langle H \rangle = -\frac{e^2}{2a}.$$

In this case this gives the exact wave function and ground state energy, but only due to the fortunate choice for the form of the wave function. In general, one would obtain an approximate wave function with on overestimate of the ground state energy.

Variational calculations are popular in a variety of many-body applications where the interactions and many-body wave functions can be extremely complicated.

6.3 The Sudden Approximation

The sudden approximation is typically applied when calculating transition probabilities for cases where the Hamiltonian changes rapidly between two times t_1 and t_2. The approximation becomes exact in the limit that the Hamiltonian changes instantaneously. In that limit one can approximate the transition amplitude by a simple overlap,

$$P(\beta \to \alpha) = |\langle \alpha | U(t_2, t_1) | \beta \rangle|^2, \tag{6.11}$$
$$\langle \alpha | U(t_2, t_1) | \beta \rangle \approx \langle \alpha | \beta \rangle,$$

where $|\beta\rangle$ is usually an eigenstate of the Hamiltonian before t_1, and $|\alpha\rangle$ is usually an eigenstate of the Hamiltonian after t_2. A previous example of a sudden transition was presented in

Example 2.1.

The sudden approximation can be used whenever the times t_1 and t_2 are so close that the relevant energy changes multiplied by $(t_2 - t_1)/\hbar$ are small. The sudden approximation is commonly used in nuclear reaction theory. If a high-energy particle knocks a proton or neutron out of the nucleus with high energy, the remainder of the nucleus finds itself suddenly in the wrong wave function, and calculating the probability it will be found in a particular state is approximated by simply taking the overlap of wave functions.

When a potential is changed slowly, the probability remains assigned to the same state. This is because for each differential change in the potential, a differential change ϵ is induced in the wave function. However, if the changes occur at much different times, as would be the case for a slow change, the differential amplitudes contribute with uncorrelated phases and the net change in the probability goes as $\sum |\epsilon|^2 = 0$. Thus, if a particle is in the ground state, and the well changes slowly, it remains in the ground state afterward. This also implies that entropy is not generated, hence the term adiabatic.

Example 6.4: Expanding Well
The most common example used to illustrate the sudden approximation is the case of a particle of mass m in an expanding well. Here, we consider an infinite square well confining particles to the region $0 < x < a$ which suddenly expands at time $t = 0$ to allow particles to occupy the region $0 < x < 2a$.
Assuming a particle was in the ground state of the old well,

a) What is the probability of being in the state n of the new well? The wave functions of the new eigenstates are

$$\psi_n(x) = \sqrt{\frac{1}{a}} \sin\left[n\pi x/(2a)\right], \quad 0 < x < 2a,$$

where the new ground state has $n = 1$.

b) What is the expectation of the energy $\langle H \rangle$ after the expansion of the well?

Solution:

a) Take the overlap of the two ground state wave functions then square it,

$$
\begin{aligned}
\boldsymbol{P(0\to n')} &= \boldsymbol{|\mathcal{M}_{0\to n}|^2},\\
\boldsymbol{\mathcal{M}_{0\to n}} &= \sqrt{\frac{2}{a}}\sqrt{\frac{1}{a}}\int_0^a \boldsymbol{dx}\ \sin(\pi x/a)\sin(n\pi x/2a)\\
&= \frac{\sqrt{2}}{4a}\int_0^a \boldsymbol{dx}\ \left[-e^{i(n+2)\pi x/2a} - e^{i(-n-2)\pi x/2a} + e^{i(n-2)\pi x/2a} + e^{i(-n+2)\pi x/2a}\right]\\
&= \frac{\sqrt{2}}{2a}\int_0^a \boldsymbol{dx}\ \{\cos[(n-2)\pi x/2a] - \cos[(n+2)\pi x/2a]\}\\
&= \frac{\sqrt{2}}{\pi}\left\{-\frac{\sin[(n+2)\pi/2]}{(n+2)} + \frac{\sin[(n-2)\pi/2]}{(n-2)}\right\}\\
&= \frac{\sqrt{2}}{\pi}\sin(n\pi/2)\left\{\frac{1}{n+2} - \frac{1}{n-2}\right\}\\
&= \begin{cases} \frac{4\sqrt{2}}{\pi}(-1)^{(n+1)/2}/(n^2-4), & n = \text{odd}\\ \frac{1}{\sqrt{2}}, & n = 2\\ 0, & n > 2 \text{ and even}\end{cases}\\
\boldsymbol{P(0\to n)} &= \begin{cases} 32/[\pi^2(n^2-4)^2], & n = \text{odd}\\ \frac{1}{2}, & n = 2\\ 0, & n > 2 \text{ and even}\end{cases}
\end{aligned}
$$

The ground state is $n = 1$ and that probability is $32/9\pi^2$.

b) The energy of the states n are $E_n = \hbar^2 n^2\pi^2/8a^2$, summing over n multiplied by $P(0\to n)$ should give the original energy because the wave function does not change in the instant the well is moved, and the kinetic energy depends only on the form of the wave function, and not on the potential. Thus, the kinetic energy will not change if the potential suddenly changes. However, if the potential changes in the region where the wave function is non-zero, the total energy will change. Thus, one could put in great effort to perform the sum but the result would be

$$\sum_n E_n|\mathcal{M}_{0\to n}|^2 = \frac{\hbar^2\pi^2}{2ma^2}.$$

6.4 Stationary-State Perturbation Theory

Here, we present stationary-state perturbation theory which is also known as Rayleigh-Schrödinger perturbation theory. Here, one solves for states and their energies in an expansion of powers of λ where the Hamiltonian is

$$H = H_0 + \lambda V. \tag{6.12}$$

The potential V is known as the perturbation and is assumed to be small while λ is assumed to be unity and is only used to keep tally of the expansion of V, i.e. the expansion in λ is an

expansion in V.

Before one embarks on perturbation theory, one assumes that one has already solved for the states $|n\rangle$, which are eigenstates of H_0, with eigenenergies ϵ_n. The goal is to express solutions for the new eigenenergies and eigenfunctions of the full Hamiltonian, H, as an expansion in powers of λ, where each term in the expansion is expressed in terms of $|n\rangle$, $\langle m|V|n\rangle$ and ϵ_n.

We assume that both the eigenstates and eigenenergies of the new Hamiltonian can be written as an expansion if powers of λ,

$$\begin{aligned} |N\rangle &= |n\rangle + \lambda|N^{(1)}\rangle + \lambda^2|N^{(2)}\rangle + \cdots \\ E_n &= \epsilon_n + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots \end{aligned} \tag{6.13}$$

Here, the terms $|N^{(j)}\rangle$ and $E_n^{(j)}$ denote the corrections to the specific eigenstate $|n\rangle$ and energies of order λ^j. We are also free to make an assumption about normalization of the state $|N\rangle$.

$$\langle n|N\rangle = 1, \tag{6.14}$$

which is equivalent to saying that the additional parts of the wave function have no $|n\rangle$ component,

$$\langle n|N^{(j)}\rangle = 0, \text{ for } j > 0. \tag{6.15}$$

The Schrödinger equation,

$$(H_0 + \lambda V)|N\rangle = E_n|N\rangle, \tag{6.16}$$

must be satisfied to every power of λ individually. By inspecting the j^{th} power of λ, this gives

$$H_0|N^{(j)}\rangle + V|N^{(j-1)}\rangle = \sum_{k=0,j} E_n^{(k)}|N^{(j-k)}\rangle. \tag{6.17}$$

Here, the sum over k goes from zero to j with the understanding that $|N^{(0)}\rangle = |n\rangle$ and $E^{(0)} = \epsilon_n$. We solve for the expressions iteratively. That is, one first finds $E^{(k)}$ then finds $|N^{(k)}\rangle$, then given those states move onto $k+1$. To find $E_n^{(k))}$, one takes the overlap of Eq. (6.17) with $\langle n|$ and using the normalization definitions one obtains

$$\langle n|V|N^{(j-1)}\rangle = E_n^{(j)}. \tag{6.18}$$

For the case where $j = 1$, one obtains the lowest-order perturbation theory answer for the energy.

$$E_n^{(1)} = \langle n|V|n\rangle. \tag{6.19}$$

The state $|N^{(j)}\rangle$ is defined by its overlap with the states $\langle m|$ where $m \neq n$. Given that $E_n^{(j)}$ is now known, one can find the $\langle m|N^{(j)}\rangle$ by taking the overlap of $\langle m|$ with Eq. (6.17),

$$\epsilon_m\langle m|N^{(j)}\rangle + \langle m|V|N^{(j-1)}\rangle = \sum_{k=0,j} E_n^{(k)}\langle m|N^{(j-k)}\rangle. \tag{6.20}$$

Solving for the j^{th} order, which is unknown,

$$\langle m|N^{(j)}\rangle = \frac{1}{\epsilon_m - \epsilon_n}\left(-\langle m|V|N^{(j-1)}\rangle + \sum_{k=1,j} E_n^{(k)}\langle m|N^{(j-k)}\rangle\right). \tag{6.21}$$

The first-order correction to the wave function is then

$$|N^{(1)}\rangle = -\sum_{m\neq n} |m\rangle \frac{1}{\epsilon_m - \epsilon_n}\langle m|V|n\rangle. \tag{6.22}$$

Using the state $|N^{(1)}\rangle$, one can then find the expression for $E_n^{(2)}$,

$$E_n^{(2)} = -\sum_{m\neq n} \frac{|\langle m|V|n\rangle|^2}{\epsilon_m - \epsilon_n}. \tag{6.23}$$

Several important principles can be realized by observing the form of $E_n^{(2)}$. First, two state's energies are pushed apart in 2nd-order perturbation theory, a phenomena known as level repulsion. Secondly, if the levels are initially close, the energies are more affected. In fact, if they are degenerate, perturbation theory breaks down, and one must apply degenerate perturbation theory which is the topic of the next sub-section. Of special significance is noticing that the ground state is always lowered in second-order perturbation theory.

Example 6.5: Harmonic Oscillator with Linear Perturbation
We will find the second-order correction to the ground state energy of the Harmonic oscillator with a linear perturbation,

$$\begin{aligned} H_0 &= -\frac{\hbar^2}{2m}\partial_x^2 + \frac{1}{2}m\omega^2 x^2, \\ V &= \beta x, \end{aligned}$$

Perhaps surprisingly, this gives the exact answer for this case.
Solution:
To proceed one can write the perturbation in terms of creation and destruction operators,

$$V = \beta\sqrt{\frac{\hbar}{2m\omega}}(a + a^\dagger).$$

The ground state then overlaps with only the first excited state,

$$\langle 1|V|0\rangle = \beta\sqrt{\frac{\hbar}{2m\omega}},$$

and the correction to the energy is

$$E^{(2)} = -\frac{|\langle 1|V|0\rangle|^2}{\hbar\omega} = -\frac{\beta^2}{2m\omega^2}.$$

One could have also solved this exactly by completing the square of the potential,

$$\begin{aligned} V(x) &= \frac{1}{2}m\omega^2 x^2 + \beta x \\ &= \frac{1}{2}m\omega^2\left(x + \frac{\beta}{m\omega^2}\right)^2 - \frac{\beta^2}{2m\omega^2}, \end{aligned}$$

and the last term is a constant correction to the energy, while the rest of the change is simply a translation of the potential. Given that the exact solution is proportional to β^2, it must also equal the perturbative correction to second order because any other order of correction would involve β to a higher power.

Finally, we point out that there is a competing scheme for perturbation theory, know as Brillouin-Wigner theory. It differs in that $E^{(n)}$ appears in the expressions for higher-order corrections to the wave functions. Operationally, results are the similar as the Rayleigh-Schrödinger form. The Brillouin-Wigner scheme will not be pursued here, but the interested reader can easily find descriptions of it in standard texts. Baym's Lectures on Quantum Mechanics contains a nice description.

6.5 Degenerate-State Perturbation Theory

Due to terms of the form

$$\frac{\langle m|V|n\rangle}{\epsilon_m - \epsilon_n},$$

perturbation theory falls apart when the perturbation mixes degenerate states, $\epsilon_m = \epsilon_n$. This can be corrected by first separating the part of the potential that mixes the degenerate states, V_d, from the remainder of the potential, V', then diagonalizing V_d and using perturbation theory for V'.

Formally, this means dividing the potential into parts using projection operators,

$$\begin{aligned} V &= V_d + V' \\ V_d &= P_d V P_d \\ V' &= (\mathbb{I} - P_d)V(\mathbb{I} - P_d) + P_d V(\mathbb{I} - P_d) + (\mathbb{I} - P_d)V P_d, \end{aligned} \tag{6.24}$$

where P_d is the projection operator that projects the subset of states that are degenerate

$$P_d = \sum_{m \in d} |m\rangle\langle m|. \tag{6.25}$$

Because $\langle m_1|V'|m_2\rangle = 0$ when m_1 and m_2 are in the degenerate set, there is no longer any problem. Note that by diagonalizing V_d, V_d essentially becomes part of H_0.

Example 6.6: Stark Effect
Here, we consider the diagonalization of V_d and forget about the perturbative part entirely. The Stark effect refers to the placement of a hydrogen atom in an electric field,

$$V = e\mathcal{E}z.$$

and are interested in calculating the splitting of the $2s$ and $2p$ levels. In this case the matrix elements of V are zero along the diagonal of the 4×4 matrix that describes V. This arises because the three p states and the s state have specific parities while the interaction is has odd parity, and an odd-parity interaction can only result in non-zero matrix elements when sandwiched between states of opposite parities. The s state can be mixed with the p states by

the interaction. Thus, there is an off-diagonal matrix element

$$\langle n = 2, \ell = 1, m_\ell = 0 | V | n = 2, \ell = 0 \rangle \neq 0.$$

All other matrix elements are zero. Using the forms for the wave functions given in a previous lecture, one can solve for the matrix element,

$$\langle 2s | V | 2p, m = 0 \rangle = \int d\Omega r^2 dr R_{2,0}(r) R_{2,1}(r) r \cos\theta Y_{0,0}(\theta, \phi) Y_{1,0}(\theta, \phi) = 3ea_0|\mathrm{E}|.$$

Because V_d, in the relevant part of the Hamiltonian matrix is a 2×2 matrix proportional to σ_x, the eigenstates of V_d are $\pm 3ea_0|\mathrm{E}|$. One can then apply perturbation theory to find corrections due to mixing with other shells. When going forward, one must use eigenstates of V_d as the basis.

6.6 The Interaction Representation (Picture)

Before one can begin time-dependent perturbation theory, we need to present the interaction representation, which is an alternative to the Schrödinger and Heisenberg representations. Summarizing the representations,

$$\begin{aligned} |\psi(t)\rangle_S &= e^{-iHt/\hbar} |\psi(t=0)\rangle_S, \\ |\psi\rangle_H &= |\psi(t=0)\rangle_S \text{ (Heisenberg states are fixed in time)}, \\ |\psi(t)\rangle_I &= e^{iH_0 t/\hbar} |\psi(t)\rangle_S = e^{iH_0 t/\hbar} e^{-iHt/\hbar} |\psi(t=0)\rangle_S, \end{aligned} \tag{6.26}$$

where the Hamiltonian is divided into an understood part, H_0, and a perturbation V. In almost all instances, the basis is chosen to correspond to eigenstates of H_0. Note that if $V = 0$ that the interaction states become equal to the Heisenberg states.

Because matrix elements $\langle \phi | ABC | \psi \rangle$ must not depend on the representation, operators in the Heisenberg and Interaction representations must be re-defined. In terms of A_S, the operator in the Schrödinger representation, then become

$$\begin{aligned} A_H(t) &= e^{iHt/\hbar} A_S e^{-iHt/\hbar} \\ A_I(t) &= e^{iH_0 t/\hbar} A_S e^{-iH_0 t/\hbar}, \end{aligned} \tag{6.27}$$

where we have assumed that H is not time dependent. Otherwise, $e^{-iHt/\hbar}$ would be replaced with an evolution operator.

Solving for the evolution of a state in the interaction representation,

$$\begin{aligned} i\hbar \frac{d}{dt} |\psi(t)\rangle_I &= i\hbar \frac{d}{dt} e^{iH_0 t/\hbar} |\psi(t)\rangle_S \\ &= -e^{iH_0 t/\hbar} H_0 |\psi(t)\rangle_S + e^{iH_0 t/\hbar} (H_0 + V) |\psi(t)\rangle_S \\ &= e^{iH_0 t/\hbar} V e^{-iH_0 t/\hbar} e^{iH_0 t/\hbar} |\psi(t)\rangle_S \\ &= V_I(t) |\psi(t)\rangle_I. \end{aligned} \tag{6.28}$$

The subscript $\boldsymbol{I}$ is omitted in most literature and the knowledge of which representation is being used is almost always left to the astute reader, who is assumed to recognize the representation by context. If the potential has an explicit time dependence, the explicit time dependence must be absorbed into $\boldsymbol{V_I(t)}$. Working in a basis of eigenstates of $\boldsymbol{H_0}$,

$$\begin{aligned} \boldsymbol{V_S(t)} &= \sum_{mn} \boldsymbol{\beta_{mn}(t)|m\rangle\langle n|}, \\ \boldsymbol{V_I(t)} &= \sum_{mn} \boldsymbol{\beta_{mn}(t) e^{i(\omega_m-\omega_n)t}|m\rangle\langle n|}, \end{aligned} \tag{6.29}$$

where the eigenstates of $\boldsymbol{H_0}$ are $\boldsymbol{\hbar\omega_n}$. In many applications the potential does not explicitly depend on time and $\boldsymbol{\beta_{mn}}$ would not depend on time.

6.7 Time-Dependent Perturbation Theory

Many problems in quantum mechanics involve time-dependent interactions. Obvious examples include spin magnetic resonance problems where the interaction explicitly varies in time as the magnetic field is oscillating. A less obvious example is a scattering problem where the incoming wave packet slowly enters the region where it feels the potential then leaves. This problem is treated by considering asymptotic momentum states with a potential that slowly turns off and on with time, rather than with a fixed potential with wave packets. In fact, nearly all perturbative scattering (e.g. Feynmann diagrams) treatments are applications of time-dependent perturbation theory even though the potential is not actually varying with time.

Using Eq. (6.28), the evolution operator in the interaction picture is defined by

$$\begin{aligned} \boldsymbol{|\psi(t)\rangle_I} &= \boldsymbol{U_I(t,t_0)|\psi(t_0)\rangle_I}, \\ &= \boldsymbol{e^{iH_0(t-t_0)/\hbar} e^{-iH(t-t_0)/\hbar}|\psi(t_0)\rangle_I}, \\ \boldsymbol{U_I(t,t_0)} &= \boldsymbol{e^{iH_0(t-t_0)/\hbar} e^{-iH(t-t_0)/\hbar}}. \end{aligned} \tag{6.30}$$

The equations of motion for the evolution operator are then,

$$\begin{aligned} \boldsymbol{i\hbar\frac{d}{dt}U(t,t_0)} &= \boldsymbol{i\hbar\frac{d}{dt}\left(e^{iH_0t/\hbar}e^{-iHt/\hbar}\right)} \\ &= \boldsymbol{e^{iH_0t/\hbar}Ve^{-iHt/\hbar}} \\ &= \boldsymbol{V_I(t)U(t,t_0)}. \end{aligned} \tag{6.31}$$

The boundary condition is

$$\boldsymbol{U(t=t_0,t_0)=1}. \tag{6.32}$$

One can rewrite the differential equation as an integral equation,

$$\boldsymbol{U_I(t,t_0) = 1 - \frac{i}{\hbar}\int_{t_0}^{t} dt' V_I(t') U_I(t',t_0)}. \tag{6.33}$$

One can see that this is a solution to the differential equation by taking derivatives of it w.r.t. time. One can take the expression for $\boldsymbol{U_I}$ on the left-hand side and insert into the place of

$\boldsymbol{U_I(t', t_0)}$ on the right-hand side. If one continues this substitution ad nauseum,

$$\boldsymbol{U_I(t,t_0) = 1 - \frac{i}{\hbar}\int_{t_0}^{t} dt' V_I(t') + \left(\frac{-i}{\hbar}\right)^2 \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' V_I(t') V_I(t'')} \tag{6.34}$$
$$\boldsymbol{\cdots + \left(\frac{-i}{\hbar}\right)^n \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' \cdots \int_{t_0}^{t^{(n-1)}} dt^{(n)} V_I(t') V_I(t'') \cdots V_I(t^{(n)}) + \cdots .}$$

The integral equation is known as the Dyson series, named after Freeman J. Dyson, https://en.wikipedia.org/wiki/Freeman_Dyson. Keeping only the first term in the expansion in $\boldsymbol{V}$ amounts to first-order perturbation theory.

Also note the the operator $\boldsymbol{U_I}$ is related to the Schrödinger evolution operator by

$$\boldsymbol{U_S = e^{-iH_0 t/\hbar} U_I e^{iH_0 t/\hbar}.} \tag{6.35}$$

Transition probabilities, which are the square of matrix elements $\boldsymbol{|\langle n|U_S|i\rangle|^2}$ with $\boldsymbol{n}$ and $\boldsymbol{i}$ referring to eigenstates of $\boldsymbol{H_0}$, are similarly $\boldsymbol{|\langle n|U_I|i\rangle|^2}$ in the interaction representation.

To second order perturbation theory, the matrix element becomes

$$\boldsymbol{\langle n|U_I(t,t_0)|i\rangle = -\frac{i}{\hbar}\int_{t_0}^{t} dt' \langle n|V_S(t')|i\rangle e^{i(\epsilon_n - \epsilon_i)t/\hbar}} \tag{6.36}$$
$$\boldsymbol{+ \left(\frac{-i}{\hbar}\right)^2 \sum_m \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' \langle n|V_S(t')|m\rangle\langle m|V_S(t'')|n\rangle e^{i(\epsilon_n - \epsilon_m)t'/\hbar} e^{i(\epsilon_m - \epsilon_i)t''/\hbar}.}$$

It is straight-forward to find expressions for the higher-order terms, but the expressions become lengthy. Note that the this expression was accomplished by invoking

$$\boldsymbol{\langle n|V_I(t)|m\rangle = \langle n|V_S(t)|m\rangle e^{i(\epsilon_n - \epsilon_m)t/\hbar},} \tag{6.37}$$

where the Schrödinger operator $\boldsymbol{V_S(t)}$ usually has no time dependence unless the interaction explicitly depends on time.

6.8 Fermi's Golden Rule

We will now derive one of the most important expressions derived in this class, Fermi's Golden Rule for the transion rate from the state $\boldsymbol{|i\rangle}$ to the state $\boldsymbol{|n\rangle}$, where $\boldsymbol{|i\rangle}$ and $\boldsymbol{|n\rangle}$ are eigenstates of $\boldsymbol{H_0}$. The transition probability will be calculated to first order in perturbation theory. We will derive the same expression twice to demonstrate how robust the derivation is.

First we consider the case where we turn on the perturbation slowly according to an exponential,

$$\boldsymbol{\langle n|V_S(t)|m\rangle = e^{\eta t} V_{nm}, \;\; \eta \to 0^+} \tag{6.38}$$

where the times being considered are less than or near zero, meaning the interaction was turned on slowly from $\boldsymbol{t = -\infty}$. To first order,

$$\boldsymbol{\langle n|U_I(t,-\infty)|i\rangle = -\frac{i}{\hbar} V_{ni} \int_{-\infty}^{t} dt' e^{i(\epsilon_n - \epsilon_i)t'/\hbar + \eta t'}} \tag{6.39}$$
$$\boldsymbol{= \frac{V_{ni}}{\epsilon_n - \epsilon_i - i\hbar\eta} e^{i(\epsilon_n - \epsilon_i)t/\hbar + \eta t}}$$

Thus, the probability goes as,

$$\begin{aligned} P_{i\to n}(t) &= |\langle n|U_I(t,-\infty)|i\rangle|^2 \\ &= \frac{|V_{ni}|^2}{(\epsilon_n-\epsilon_i)^2+\hbar^2\eta^2}e^{2\eta t}, \end{aligned} \tag{6.40}$$

In obtain Fermi's Golden Rule, the goal is to find the transition rate,

$$\begin{aligned} R_{i\to n} &= \frac{d}{dt}P_{i\to n}(t) \\ &= \frac{2\eta|V_{ni}|^2}{(\epsilon_n-\epsilon_i)^2+\hbar^2\eta^2}, \end{aligned} \tag{6.41}$$

where we have made the approximation that we are considering the rate at a time where ηt is small. From an end-of-chapter problem, you can see that the η-dependent terms can be replaced with a δ function as $\eta \to 0^+$,

$$R_{i\to n}(t) = \frac{2\pi}{\hbar}|V_{ni}|^2\delta(\epsilon_n-\epsilon_i) \tag{6.42}$$

Thus, there is only a non-zero transition rate if the final state has the same energy as the initial state. This works for scattering or decays, where indeed there are many final states with a given energy. The δ function looks a bit peculiar, but makes sense when the states n are in a continuum. For instance, one does not calculate the electromagnetic decay rate of a radioactive nucleus to a specific state where a photon has a momentum k, but instead to any state within a specific angle. By summing over all such states, and using the density of states to transform the sum to an integral over the photon energy, $\int d\epsilon_\gamma \cdots$, the delta function is effectively replaced by the density of photon states.

Aside: Changing Sum over Momentum States to Integral

Fermi's golden rule features a delta function in energy, $\delta(E_f - E_i)$. This suggests you can only transform to states with exactly the same energy, and if the state does have exactly the same energy, the transition rate would explode. In reality, the delta function has a finite width, $\hbar\eta$, and only as $\eta \to 0$, $\delta(E_f - E_i)$ becomes a true delta function. Another limit one considers is that momentum states are defined in some length, or some volume in three dimensions, that approaches infinity. Summing over final states means summing over final momenta, which in the limit of $L \to \infty$ means that the number of states within some small energy range will become infinity. Thus, we will assume that $\eta \to 0$, but that the length approaches ∞ such that $\eta L \to \infty$.

As shown in Chapter 2, if one has a one-dimensional infinite square well of length L, the boundary conditions give

$$\begin{aligned} \sin(kL) &= 0, \\ kL &= n\pi, \quad (k > 0) \\ \frac{dn}{dk} &= \frac{L}{\pi}, \end{aligned} \tag{6.43}$$

If one considers momentum states, rather than the $\sin(kx)$ forms which describe a linear combination of both forward and backward waves, the number of states per dk is

$$\frac{dn}{dk} = \frac{L}{(2\pi)}, \quad -\infty < k < \infty. \tag{6.44}$$

One can extend this into three dimensions where the infinite volume is Ω,

$$\frac{d^3n}{d^3k} = \frac{\Omega}{(2\pi)^3}. \tag{6.45}$$

If the particle also had spin, a factor of $(2s+1)$ would be added. The sum over states in three dimensions thus becomes

$$\begin{aligned} \sum_k &\to \frac{L}{2\pi}\int_{-\infty}^{\infty} dk, \ \text{one dimension}, \\ \sum_k &\to \frac{A}{(2\pi)^2}\int_{-\infty}^{\infty} d^2k, \ \text{two dimensions}, \\ \sum_k &\to \frac{\Omega}{(2\pi)^3}\int d^3k, \ \text{three dimensions}. \end{aligned} \tag{6.46}$$

For Fermi's golden rule, the delta function is expressed in energy, so in order to integrate the differential dk must be changed to dE,

$$\begin{aligned} dk &= \frac{dE}{dE/dk} = \frac{dE}{\hbar v_k}, \\ &= dE\frac{m}{\hbar^2 k}, \ \text{non}-\text{relativistic}, \\ &= dE\frac{E}{\hbar^2 k}, \ \text{relativistic}, \\ &= \frac{dE}{\hbar c}, \ \text{massless}. \end{aligned} \tag{6.47}$$

Here, $v_k = dE/dp$ is the velocity of a particle of momentum p. Integrals of the type one encounters with Fermi's Golden rule can then be easily performed,

$$\int dk\, f(k)\delta(E_f - E_i) = \left.\frac{f(k)}{\hbar v_k}\right|_{\text{with } k \text{ such that } E_f = E_i}. \tag{6.48}$$

6.9 Harmonic Perturbations

Rather than having a potential that turned on slowly, one could envisage a potential with an explicit time dependence characterized by some frequency ω,

$$\langle n|V_S(t)|m\rangle = V_{nm}e^{\eta t}\cos(\omega t) = \frac{1}{2}V_{nm}e^{\eta t}\left(e^{i\omega t}+e^{-i\omega t}\right). \tag{6.49}$$

One can then follow the same derivation by noting that the only two differences are the factor of $1/2$, which get squared to obtain a probability, and the extra phases in the two terms. If we consider each term separately $e^{i\omega t}$ and $e^{-i\omega t}$, one can write the answer without much thought.

$$\frac{d}{dt}P_{i\to n}(t) = \frac{2\pi}{\hbar}\frac{|V_{ni}|^2}{4}\left[\delta(\epsilon_n-\epsilon_i+\hbar\omega)+\delta(\epsilon_n-\epsilon_i-\hbar\omega)\right]. \tag{6.50}$$

Additionally there would be cross term from squaring the matrix elements, but that term behaves as $\cos(2\omega t)$ and is disregarded when averaging over time.

Example 6.7: Exciting a Particle with a Radiative-Like Interaction
Consider a particle of mass m in the ground state of a δ function potential,

$$V_0(x) = -\beta\delta(x)$$

The particle feels a harmonic potential

$$V(t) = eEx\cos(\omega t),\ \ \hbar\omega > |\text{G.S. energy}|$$

Estimate the ionization rate using first-order perturbation theory. To simplify the problem, assume the outgoing momentum is high enough that the outgoing wave can be treated as a plane wave, i.e. the corrections due to the delta function potential are small. This is a one-dimensional example that has much in common with radiative excitation.
To solve this problem,

1. Find the G.S. wave function and G.S energy of the delta function potential.
2. Calculate the matrix element $\langle k|V|0\rangle$ using a normalized state $\langle x|k\rangle = e^{ikx}/\sqrt{L}$ where L is an arbitrarily large size to the box.
3. Sum the probability over states by writing it as a density of states. Notice how the size of the box falls out of the problem.

(1) The solution to the delta function potential from Example (2.4), and the plane wave functions are,

$$\psi_0(x) = \sqrt{q/2}e^{-q|x|},\ q = \frac{m\beta}{\hbar^2},$$
$$\psi_k(x) = \frac{e^{ikx}}{\sqrt{L}},\ k = \sqrt{2m(\hbar\omega - B)/\hbar^2},$$

where $B = \hbar^2 q^2/2m$ is the binding energy.
(2) The matrix element is

$$\begin{aligned} V_{k0} &= \sqrt{\frac{q}{2L}} \int dx\, eExe^{-ikx} e^{-q|x|} \\ &= -ieE\sqrt{\frac{q}{2L}} \int dx\, x \sin(kx) e^{-q|x|} \\ &= ieE\sqrt{\frac{q}{2L}} \frac{d}{dk} \int dx\, \cos(kx) e^{-q|x|} \\ &= ieE\sqrt{\frac{q}{2L}} \frac{d}{dk} \left\{ \frac{1}{q+ik} + \frac{1}{q-ik} \right\} \\ &= ieE\sqrt{\frac{2}{L}} \frac{2kq^{3/2}}{(q^2+k^2)^2}. \end{aligned}$$

(3) Using Fermi's golden rule, the ionization rate is

$$\begin{aligned} \mathcal{R} &= \frac{\pi}{2\hbar} \int \frac{Ldk}{2\pi} |V_{k0}|^2 \delta(\epsilon_k - \hbar\omega - B) \\ &= \frac{1}{\hbar} \frac{e^2E^2k^2q^3}{(q^2+k^2)^4} \frac{4}{|d\epsilon_k/dk|} \\ &= \frac{4m}{\hbar^3} \frac{e^2E^2kq^3}{(q^2+k^2)^4}. \end{aligned}$$

A factor of two entered the expression because the delta function picked up contributions from $\pm k$.
As stated in the problem, an approximation was made in that the outgoing wave function was assigned the form $e^{ikx}/\sqrt{L}$. More realistically, the outgoing wave should have been a plane wave modified by the delta function potential. In that case, the form would be

$$\psi_k(x) = \frac{1}{\sqrt{L}} \begin{cases} e^{ikx} + Ae^{-ikx}, & x > 0 \\ Be^{ikx}, & x < 0 \end{cases}.$$

A similar form would have been used for the solution with outgoing wave e^{-ikx}. One could have solved for A and B, then calculated the integrals as done before. This improvement requires some work, but is certainly tractable. In nuclear physics fixing the outgoing wave to include the potential is known as the distorted wave Born approximation (DWBA).

6.10 An Exactly Solvable Model of a Time-Dependent Interaction

The problem described here is quite specific, as it only applies to two-component problems. However, it does illustrate how time dependence manifests itself in the evolution. Consider two

states $|\mathbf{1}\rangle$ and $|\mathbf{2}\rangle$ interacting through the time-dependent potential,

$$\begin{aligned}\boldsymbol{H_0} &= \hbar\boldsymbol{\omega_1}|\mathbf{1}\rangle\langle\mathbf{1}| + \hbar\boldsymbol{\omega_2}|\mathbf{2}\rangle\langle\mathbf{2}| = \frac{\hbar}{\mathbf{2}}(\boldsymbol{\omega_1}+\boldsymbol{\omega_2})\mathbb{I} + \frac{\hbar}{\mathbf{2}}(\boldsymbol{\omega_1}-\boldsymbol{\omega_2})\boldsymbol{\sigma_z}\\ \boldsymbol{V_t} &= \boldsymbol{\gamma}\cos(\boldsymbol{\omega t})\boldsymbol{\sigma_x} + \boldsymbol{\gamma}\sin(\boldsymbol{\omega t})\boldsymbol{\sigma_y}\end{aligned}$$

or written as a matrix,

$$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{E_1} & \boldsymbol{\gamma e^{-i\omega t}} \\ \boldsymbol{\gamma e^{i\omega t}} & \boldsymbol{E_2} \end{pmatrix}. \tag{6.51}$$

This is the form of an interaction with a time-dependent magnetic field,

$$\boldsymbol{H} = \boldsymbol{H_0} + \vec{\boldsymbol{\mu}}\cdot\vec{\boldsymbol{B}}(\boldsymbol{t}),\ \ \boldsymbol{B}(\boldsymbol{t}) = \boldsymbol{B_0}\hat{\boldsymbol{z}} + \boldsymbol{B_\perp}\left[\hat{\boldsymbol{x}}\cos(\boldsymbol{\omega t}) + \hat{\boldsymbol{y}}\sin(\boldsymbol{\omega t})\right]. \tag{6.52}$$

One can make the substitution,

$$|\psi\rangle \rightarrow \boldsymbol{e^{-i\bar{\omega}t}}|\psi\rangle,\ \bar{\boldsymbol{\omega}} \equiv \frac{\boldsymbol{E_1}+\boldsymbol{E_2}}{\mathbf{2}\hbar} \tag{6.53}$$

to write the new Hamiltonian as

$$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{\omega_{12}/2} & \boldsymbol{\gamma e^{-i\omega t}} \\ \boldsymbol{\gamma e^{i\omega t}} & \boldsymbol{-\omega_{12}/2} \end{pmatrix}, \tag{6.54}$$

where $\boldsymbol{\omega_{12}} \equiv (\boldsymbol{E_1}-\boldsymbol{E_2})/\mathbf{2}$.
Writing the evolution in the interaction picture for the components $\boldsymbol{\psi_1}$ and $\boldsymbol{\psi_2}$,

$$\begin{aligned}\frac{\boldsymbol{d}}{\boldsymbol{dt}}\boldsymbol{\psi_1}(\boldsymbol{t}) &= \boldsymbol{i}\frac{\boldsymbol{\omega_{12}}}{\mathbf{2}}\boldsymbol{\psi_1}(\boldsymbol{t}) + \boldsymbol{i}\frac{\boldsymbol{\gamma}}{\hbar}\boldsymbol{e^{-i\omega t}}\boldsymbol{\psi_2}(\boldsymbol{t})\\ \frac{\boldsymbol{d}}{\boldsymbol{dt}}\boldsymbol{\psi_2}(\boldsymbol{t}) &= \boldsymbol{-i}\frac{\boldsymbol{\omega_{12}}}{\mathbf{2}}\boldsymbol{\psi_2}(\boldsymbol{t}) + \boldsymbol{i}\frac{\boldsymbol{\gamma}}{\hbar}\boldsymbol{e^{i\omega t}}\boldsymbol{\psi_1}(\boldsymbol{t})\end{aligned}$$

Now, by making the substitution,

$$\begin{aligned}\boldsymbol{\psi_1'} &\equiv \boldsymbol{e^{i\omega t/2}}\boldsymbol{\psi_1}\\ \boldsymbol{\psi_2'} &\equiv \boldsymbol{e^{-i\omega t/2}}\boldsymbol{\psi_2},\end{aligned}$$

one can derive the evolutions,

$$\begin{aligned}\frac{\boldsymbol{d}}{\boldsymbol{dt}}\boldsymbol{\psi_1'}(\boldsymbol{t}) &= \boldsymbol{-i}\frac{(\boldsymbol{\omega_{12}}-\boldsymbol{\omega})}{\mathbf{2}}\boldsymbol{\psi_1'}(\boldsymbol{t}) + \frac{\boldsymbol{\gamma}}{\hbar}\boldsymbol{\psi_2'}(\boldsymbol{t})\\ \frac{\boldsymbol{d}}{\boldsymbol{dt}}\boldsymbol{\psi_2'}(\boldsymbol{t}) &= \boldsymbol{i}\frac{(\boldsymbol{\omega_{12}}-\boldsymbol{\omega})}{\mathbf{2}}\boldsymbol{\psi_2'}(\boldsymbol{t}) + \frac{\boldsymbol{\gamma}}{\hbar}\boldsymbol{\psi_1'}(\boldsymbol{t}).\end{aligned}$$

The problem now looks like a Hamiltonian without a time dependent interaction,

$$\boldsymbol{H'} = \frac{\hbar(\boldsymbol{\omega_{12}}-\boldsymbol{\omega})}{\mathbf{2}}\boldsymbol{\sigma_z} + \boldsymbol{\gamma}\boldsymbol{\sigma_y}. \tag{6.55}$$

This is the same problem as we worked out for neutrino oscillations. The evolution operator becomes

$$\begin{aligned} e^{-iH't/\hbar} &= \cos(\Omega t) + \sigma_n \sin(\Omega t), \\ \Omega &= \sqrt{\frac{(\omega_{12}-\omega)^2}{4} + \frac{\gamma^2}{\hbar^2}} \\ \sigma_n &= \cos(\theta)\sigma_z + \sin(\theta)\sigma_y \\ \tan(\theta) &= \frac{2\gamma}{\hbar(\omega_{12}-\omega)} \end{aligned}$$

If the state begins polarized in the z direction, the maximum probability of becoming a spin-down state is

$$\max |\langle \downarrow | e^{-iH't/\hbar} | \uparrow \rangle|^2 = \frac{\gamma^2/\hbar^2}{\gamma^2/\hbar^2 + (\omega-\omega_{12})^2/4} \tag{6.56}$$

The resonant form in this example clearly displays that driving the system at the resonant frequency, $\omega = \omega_{12}$, results in the greatest chance for flipping the spin. NMR works on very similar principles, only in this case the time-dependent field usually oscillates only in one plane, i.e.

$$\vec{B}(t) = B_0 \hat{z} + B_\perp \hat{x} \cos(\omega t).$$

This is a bit harder to work out as compared to our example but the resonant conditions remain the same.

The functional form above,

$$f(\omega) = \frac{1}{\pi} \frac{\Gamma/2}{1 + (\omega-\omega_r)^2/(\Gamma/2)^2}, \tag{6.57}$$

is known as a Lorentzian. This function has a maximum at $\omega = \omega_R$, and integrates to unity. The parameter Γ is known as the width, or the half-width, meaning that the width of the function measured at half its maximum height is Γ. Lorentzian forms show up in resonant forms, e.g. for driven harmonic oscillators. When the driving frequency, ω, matches the fundamental, or resonant, frequency, ω_R, the resulting amplitude is a maximum. In quantum mechanics, adjusting the energy can be thought of as adjusting the frequency, and when an energy from some channel matches the energy of a different channel, the channels couple most strongly. For instance, if the beam energy of scattering matches the energy of an unstable state, the scattering cross sections often maximize at that energy. For this reason, excited states encountered in nuclear or particle physics are often called resonances.

6.11 Exercises

1. Using the WKB approximation, you can estimate the lifetime of a particle of mass m initially trapped in the "ground state" of a one-dimensional rectangular well using Eq. (6.10),

$$V(x) = \begin{cases} \infty, & x < 0 \\ 0, & 0 < x < a \\ \frac{\alpha}{x}, & a < x \end{cases}$$

Assume the barrier is sufficiently high that the wave function in the well can be approximated as that of an infinite well and that the frequency of tunneling attempts can be thought of as the rate at which a classical particle would impact the barrier at that energy.

(a) Estimate the energy E, by treating the potential as if its infinity at $x = a$.
(b) To what point, x_f, does the particle need to tunnel to escape the well.
(c) Estimate the tunneling probability, $e^{-2\phi}$, where ϕ is calculated using the WKB approximation. I.e. find an expression for ϕ. Warning: The integral is not trivial.
(d) In terms of ϕ what is the lifetime?

2. A particle of mass m is initially in the ground state of a one-dimensional harmonic oscillator of frequency ω centered at $x = 0$. Suddenly, at time $t = 0$, the center of the well is moved to $x = \ell$. Here, you will calculate the probability that the particle will be found in the state $|n\rangle$ of the new well, where $|0\rangle$ is the new ground state.

(a) The ground state of the old well can be written as

$$|\phi_0\rangle = e^{\ell(d/dx)}|0\rangle,$$

where ϕ_0 is the ground state of the old well and $|n\rangle$ refers to eigenstates of the newly positioned well. Writing the operator d/dx in terms of creation and destruction operators of the new well, find an expression for $|\phi_0\rangle$ as a linear combination of $|n\rangle$. HINT: You will probably wish to use the Baker-Campbell-Hausdorff relation.
(b) What is the probability of finding the particle in the state $|n\rangle$.
(c) What is the expectation of the energy $\langle H\rangle$ after the well is shifted?
(d) If the well were shifted slowly instead of suddenly to its new position, what would be the probability of finding the particle in the ground state of the new well?

3. Estimate the ground state energy of the hydrogen atom using a three-dimensional harmonic oscillator ground state wave function as a trial function. Give your answer in terms of the electron mass m and the coupling e^2.

4. Estimate the ground state energy of the three-dimensional harmonic oscillator using the hydrogen atom wave function as the trial wave function. Give your answer in terms of the mass m and the characteristic frequency of the harmonic oscillator ω.

5. Consider a particle in an infinitely deep square well of width a.

$$V_0(x) = \begin{cases} \infty, & x < -a/2 \\ 0, & -a/2 < x < a/2 \\ \infty, & x > a/2 \end{cases}$$

A particle feels a perturbative potential,

$$V_1(x) = \beta \sin(\pi x/a)$$

(a) What is the change in the ground state energy in lowest (non-zero) order perturbation theory?
(b) What is the correction to the energy of the first excited state to the same order?

(c) What is the correction to the wave function of the ground state to lowest non-zero order?

6. Consider the Hamiltonian:

$$\boldsymbol{H_0 = \alpha \sigma_z}$$

and the perturbation

$$\boldsymbol{V = \beta \sigma_x}$$

(a) What is the correction to the ground state energy to second order in perturbation theory?
(b) What is the correction to the excited state's energy to the same order?
(c) Find the exact expression for the energy of the first state, and show that it gives the same answer as part $\boldsymbol{a}$ when expanded in powers of $\boldsymbol{\beta}$.

7. An electron of mass $\boldsymbol{m}$ initially in the ground state of a three-dimensional harmonic oscillator potential characterized by frequency $\boldsymbol{\omega}$, i.e. the ground state energy is $\boldsymbol{3\hbar\omega/2}$, is placed in a region with uniform electric field $\boldsymbol{E}$.

(a) By finding corrections to the ground state wave function in first order perturbation theory, find an expression for the electric dipole moment induced in the atom.
(b) An alternative method for calculating the dipole moment is to differentiate the energy with respect to the electric field. Show that this method yields the same expression found in (a) when one uses second order perturbation theory to find the correction to the energy.

8. Two electrons whose positions are defined by $\mathbf{r_1}$ and $\mathbf{r_2}$ relative to the centers of their confining potentials. The confining potentials are then separated by a distance $\vec{\boldsymbol{R}}$.

$$\boldsymbol{V_0(\mathbf{r_1}, \mathbf{r_2}) = \frac{1}{2} m\omega^2 (r_1^2 + r_2^2).}$$

Positive charges $+\boldsymbol{e}$ are fixed at the centers of the potentials. The electromagnetic energy between the two wells is:

$$\boldsymbol{V = \frac{e^2}{R} + \frac{e^2}{|\vec{R} + \vec{r}_1 - \vec{r}_2|} - \frac{e^2}{|\vec{R} + \vec{r}_1|} - \frac{e^2}{|\vec{R} - \vec{r}_2|}.}$$

Here, the repulsive interaction between the two positive ions is described by the first term, and the repulsive interaction between the electrons is described by the second term. The last two terms describe the attractive interaction between the electron and the ion in the other well. The electromagnetic energy between each electron and its confining ion is assumed to be part of the confining potential, and not part of the perturbation.

Assume that the separation $\boldsymbol{R}$ is much larger than either $\boldsymbol{r_1}$ or $\boldsymbol{r_2}$. In terms of the separation between the wells, $\boldsymbol{R}$, the mass of the electrons $\boldsymbol{m}$, the charge $\boldsymbol{e}$ and $\boldsymbol{\omega}$,

(a) Show that for large R, the interaction may be approximated as a dipole-dipole interaction,

$$V = \frac{e^2}{R^3}\left(x_1x_2 + y_1y_2 - 2z_1z_2\right),$$

where the z axis is along the direction of $\vec{R}$.

(b) Use second-order perturbation theory to find the electromagnetic attraction of the two wells, $V(R)$. This motivates the form for the London dispersion force, https://en.wikipedia.org/wiki/London_dispersion_force, which is the long-range attractive force between neutral molecules.

9. Consider the function

$$g(\omega) \equiv \mathrm{Im}\frac{1}{\omega - i\eta} = \frac{\eta}{\omega^2 + \eta^2}$$

where η is a positive real constant that approaches zero.

(a) What is $g(\omega = 0)$?

(b) What is $g(\omega \neq 0)$?

(c) Using trigonometric substitutions, evaluate

$$\int_{-\infty}^{\infty} d\omega\, g(\omega).$$

(d) Write an expression for a delta function in terms of $g(\omega)$.

10. A ***bob*** particle is in the ground state of a 3-dimensional harmonic oscillator characterized by a frequency ω,

$$V_0 = \frac{1}{2}m\omega^2 r^2$$

A perturbation is added that allows a ***bob*** particle to undergo a transformation into a ***mary*** particle. The ***mary*** particle does not feel the effects of the oscillator potential. The ***bob*** and ***mary*** particles have the same mass m. The perturbation is of the form,

$$V_{bm} = g\vec{\epsilon}_s \cdot \int d^3r \psi^*_{bob}(\mathrm{r})\nabla\psi_{mary}(\mathrm{r}),$$

where ϵ_s is the unit polarization vector of the ***mary*** particle with polarization s, which refers to any of three directions.

(a) Find the total decay rate Γ (to any polarization). Hint: The answer will of the form

$$\Gamma = A\frac{mg^2b^3k^3}{\hbar^3}, \tag{6.58}$$

where A is a dimensionless constant and b is the size of the harmonic oscillator wave function. You need to find A.

(b) Find the differential decay rate, $d\Gamma_s/d\cos\theta$, for a polarization $\vec{\epsilon}_s = \hat{z}$.

7 Perturbative Approaches to Scattering

7.1 Cross Sections

Scattering experiments provide much of the basis from which knowledge of atomic and subatomic physics is extracted. This ranges from investigating the forces between the projectile and target to understanding the elementary constituents and substructure of matter. Experiments often involve aiming a beam of projectiles at a target with a given energy, then observing the probability with which they scatter at given angles. Scattering cross sections, σ_{if}, can be related to the rate for scattering of a single target particle. This rate can then be expressed in terms of Fermi's golden rule. For an incoming wave packet with an area of extent A and a length L, the probability of colliding is

$$P_{\text{scatt}}^{(N)} = \frac{N\sigma}{A}, \tag{7.1}$$

where there are N target particles. If a target has thickness Z, area A and a number density of scatterers, n_t,

$$\begin{aligned} N &= n_t A Z, \\ P_{\text{scatt}}^{(N)} &= (n_t Z)\sigma, \\ \sigma &= \frac{P_{\text{scatt}}}{n_t Z}. \end{aligned} \tag{7.2}$$

Thus, armed with knowledge of the target density and thickness, an experimentalist can determine the cross section for a single scatterer.

The choice of using the cross section to describe the result of the scattering is to use an observable that is independent of the size and density of the target, and encapsulates the physics of scattering off a single target particle.

The cross section can be related to a rate for the scattering by a single particle by considering an incoming wave packet of longitudinal extent (along the direction of the incoming wave) L and transverse size A. Relating σ to the rate is necessary because perturbative approaches, like Fermi's golden rule, provide expressions for the rate.The rate for scattering off a single scatterer is

$$R = \frac{P_{\text{scatt}}^{(1)}}{L/v_p}, \tag{7.3}$$

where the velocity of the projectile is v_p and the time for the packet to pass is L/v_p. Given that the probability for scattering off a single scatterer is σ/A,

$$\begin{aligned} R &= v_p \frac{\sigma}{W}, \\ W &= AL. \end{aligned} \tag{7.4}$$

Here W is the volume of the wave packet. To calculate the cross section in perturbation theory one can find an expression for the rate in terms of Fermi's golden rule, then obtain the cross

section,

$$\boldsymbol{\sigma = \frac{W}{v_p} R.} \tag{7.5}$$

The dependence on the wave-packet volume, $\boldsymbol{W}$, will disappear for the theoretical expressions used to calculate the cross section, as the cross section should depend solely on the projectile mass, $\boldsymbol{m}$, the projectile velocity $\boldsymbol{v_p}$ and the potential, $\boldsymbol{V(r)}$, experienced by the particle.

Theoretical descriptions of scattering mainly fall into two classes. The first is perturbative, the simplest being the Born approximation. These are typically performed in a plane-wave basis, with the matrix elements being expressed in the form $\boldsymbol{V_{\vec{k},\vec{k}'}}$. The second class of treatments are built around a partial wave basis and are non-perturbative. Here, one considers incoming and outgoing spherical waves, known as partial waves. Using angular momentum conservation, one can divide the problem into a few one-dimensional problems, which can be solved non-perturbatively, often by solving Schrödinger's equation numerically. For lower energies, the latter becomes more tractable because only a few partial waves contribute. The momentum multiplied by the range of the potential sets the scale for what values of angular momentum are relevant. Furthermore, at low energies, the potential energy is not much smaller than the kinetic energy, which would invalidate the perturbative class of treatments, which are perturbative in $\boldsymbol{V}$. For the next several sections these notes focus on the perturbative picture, then switch to partial wave treatments in the next chapter.

If one is interested in only those scatterings from the initial momentum state $\boldsymbol{\vec{k}_i}$ to a specific final state $\boldsymbol{\vec{k}_f}$, the cross section becomes

$$\boldsymbol{\sigma_{if} = \frac{W}{v_p} R(\vec{k}_i \to \vec{k}_f).} \tag{7.6}$$

If one considers the probability per solid angle that scattered particles are emitted in a given direction, int into a solid angle element of size $\boldsymbol{d\Omega}$, the expression becomes

$$\boldsymbol{\frac{d\sigma}{d\Omega} = \frac{1}{n_t Z} \frac{dP_{scatt}}{d\Omega}.} \tag{7.7}$$

The differential solid angle is $\boldsymbol{d\Omega = \sin\theta d\theta d\phi = d\cos\theta d\phi}$, with $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ referencing the direction of the final momentum. The solid angle is a measure of the angular extent in spherical coordinates. For example, the area on a sphere is $\boldsymbol{r^2 d\Omega}$. The total angular coverage is $\boldsymbol{\int d\Omega = 4\pi}$, which is why many hermetic detectors are called $\boldsymbol{4\pi}$ detectors. The quantity $\boldsymbol{d\sigma/d\Omega}$ is referred to as the differential cross section, and integrates to the total cross section,

$$\boldsymbol{\int d\Omega \frac{d\sigma}{d\Omega} = \sigma_{tot}.} \tag{7.8}$$

In essence, experiments measure $\boldsymbol{dP_{scatt}/d\Omega}$, and given their knowledge of the thickness of the target and the density of scatterers, infer the cross section. Rather than comparing scattering rates, which depend on the size of the projectile's wave packet, the target thickness and the target's number density, experiments report on differential or total cross sections.

7.2 The Born approximation

From Eq. (7.6) one can write the cross section for scattering into any state in terms of a rate. Then, using Fermi's golden rule to approximate the rate,

$$\begin{aligned}\sigma &= \frac{W}{v}\sum_{\vec{k}_f} R(\vec{k}_i \to \vec{k}_f) \\ &= \frac{2\pi W}{v\hbar}\sum_{\vec{k}_f} |\langle \vec{k}_f|\mathcal{V}|\vec{k}_i\rangle|^2 \delta(\epsilon_f - \epsilon_i).\end{aligned} \tag{7.9}$$

One may express the sum over states as an integral,

$$\sum_{\vec{k}_f} = \frac{W}{(2\pi)^3}\int d^3k_f = \frac{W}{(2\pi)^3}\int d\Omega \frac{k^2}{\hbar v} dE, \tag{7.10}$$

and write the matrix element as,

$$\langle \vec{k}_f|\mathcal{V}|\vec{k}_i\rangle = \int d^3r \frac{e^{-i\vec{k}_f\cdot\vec{r}}}{\sqrt{W}}\mathcal{V}(\vec{r})\frac{e^{i\vec{k}_i\cdot\vec{r}}}{\sqrt{W}} \tag{7.11}$$

to obtain an expression for the cross section where the volume has canceled,

$$\begin{aligned}\sigma &= \frac{1}{4\pi^2\hbar v}\int d^3k_f\ \delta(\epsilon_f - \epsilon_i)\left|\int d^3r\ \mathcal{V}(\vec{r})e^{i(\vec{k}_f - \vec{k}_i)\cdot\vec{r}}\right|^2 \\ &= \frac{m}{4\pi^2\hbar^4 p}\int d^3p_f\ \delta(\epsilon_f - \epsilon_i)\left|\int d^3r\ \mathcal{V}(\vec{r})e^{i(\vec{k}_f - \vec{k}_i)\cdot\vec{r}}\right|^2 \\ &= \frac{m}{4\pi^2\hbar^4 p}\int pmd\epsilon_f d\Omega\ \delta(\epsilon_f - \epsilon_i)\left|\int d^3r\ \mathcal{V}(\vec{r})e^{i(\vec{k}_f - \vec{k}_i)\cdot\vec{r}}\right|^2 \\ &= \frac{m^2}{4\pi^2\hbar^4}\int d\Omega\ \left|\int d^3r\ \mathcal{V}(\vec{r})e^{i(\vec{k}_f - \vec{k}_i)\cdot\vec{r}}\right|^2.\end{aligned} \tag{7.12}$$

As promised, there is no dependence on the wave-packet volume W. This is known as the Born approximation for the cross section, and is an example of first order perturbation theory, where in this case one is calculating the rate from an initial to some final states $\vec{k}_f$ using Fermi's golden rule. To change this to the differential cross section, one can simply divide both sides by $d\Omega$ to find the differential form of the Born approximation,

$$\begin{aligned}\frac{d\sigma}{d\Omega} &= \frac{m^2}{4\pi^2\hbar^4}|\tilde{\mathcal{V}}(\vec{q})|^2, \\ \tilde{\mathcal{V}}(\vec{q}) &= \int d^3r\ \mathcal{V}(\vec{r})e^{i\vec{q}\cdot\vec{r}}, \\ \vec{q} &= \vec{k}_i - \vec{k}_f.\end{aligned} \tag{7.13}$$

The differential cross section for scattering into the direction $\vec{k}_f$ is determined by the Fourier transform of the potential, where the momentum transferred to the target, $\vec{k}_i - \vec{k}_f$, enters the

Fourier transform. For elastic scattering, where no energy is lost by the beam and $|\boldsymbol{k_i}| = |\boldsymbol{k_f}|$, the momentum transfer is related to the scattering direction noted by $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$,

$$\vec{\boldsymbol{k}}_{\boldsymbol{i}} - \vec{\boldsymbol{k}}_{\boldsymbol{f}} = |\boldsymbol{k}| \left[(\mathbf{1} - \cos\boldsymbol{\theta})\hat{\boldsymbol{z}} - \sin\boldsymbol{\theta}\cos\boldsymbol{\phi}\hat{\boldsymbol{x}} - \sin\boldsymbol{\theta}\sin\boldsymbol{\phi}\hat{\boldsymbol{y}}\right], \tag{7.14}$$

where the initial beam was traveling in the $+\hat{z}$ direction.

Example 7.1: Born Approximation for a Spherically Symmetric Gaussian Potential
Consider the potential

$$\boldsymbol{\mathcal{V}(r) = V_0 e^{-r^2/(2a^2)}.}$$

a. As a function of the scattering angle $\boldsymbol{\theta}$ find the differential cross section.

Solution:
First calculate the Fourier transform of the potential which depends only on the magnitude of $\vec{\boldsymbol{k}}_{\boldsymbol{i}} - \vec{\boldsymbol{k}}_{\boldsymbol{f}}$.

$$\begin{aligned}
\tilde{\boldsymbol{\mathcal{V}}}(\vec{\boldsymbol{q}}) &= \boldsymbol{V_0} \int \boldsymbol{d^3 r}\; \boldsymbol{e^{-r^2/(2a^2)} e^{i\vec{q}\cdot\vec{r}}} \\
&= \boldsymbol{V_0} \int \boldsymbol{d^3 r}\; \boldsymbol{e^{-(\vec{r} - i\vec{q}a^2)^2/(2a^2)} e^{-q^2a^2/2}} \\
&= \boldsymbol{V_0 a^3 (2\pi)^{3/2} e^{-q^2a^2/2}},
\end{aligned}$$

where

$$\begin{aligned}
\boldsymbol{q} &= |\vec{\boldsymbol{k}}_{\boldsymbol{i}} - \vec{\boldsymbol{k}}_{\boldsymbol{f}}| \\
&= \boldsymbol{k}\sqrt{(\mathbf{1} - \cos\boldsymbol{\theta})^2 + \sin^2\boldsymbol{\theta}} \\
&= \boldsymbol{k}\sqrt{\mathbf{2}(\mathbf{1} - \cos\boldsymbol{\theta})} \\
&= \boldsymbol{2k}\sin(\boldsymbol{\theta}/\mathbf{2}).
\end{aligned} \tag{7.15}$$

Thus, the differential cross section is

$$\frac{\boldsymbol{d\sigma}}{\boldsymbol{d\Omega}} = \frac{\boldsymbol{2\pi m^2 a^6 V_0^2}}{\hbar^4} \boldsymbol{e^{-4k^2a^2\sin^2(\theta/2)}}.$$

b. Find the total cross section in this approximation.

Solution:
This involves integrating $\boldsymbol{d\sigma/d\Omega}$ over $\boldsymbol{d\Omega = 2\pi d\cos\theta}$.

$$\begin{aligned}
\boldsymbol{\sigma} &= \int \boldsymbol{d\sigma} = \frac{\boldsymbol{4\pi^2 m^2 a^6 V_0^2}}{\hbar^4} \int \boldsymbol{e^{-4k^2a^2\sin^2(\theta/2)} d\cos\theta}, \\
&= \frac{\boldsymbol{8\pi^2 m^2 a^6 V_0^2}}{\hbar^4} \int_{\mathbf{0}}^{\mathbf{1}} \boldsymbol{d\sin(\theta/2)\; e^{4k^2a^2\sin^2(\theta/2)} \sin(\theta/2)},
\end{aligned}$$

where a double angle formula, $\cos\theta = 1 - 2\sin^2(\theta/2)$, was applied. Integrating over $\sin\theta/2$,

$$\sigma = \frac{\pi^2 m^2 a^4 V_0^2}{\hbar^4 k^2}.$$

The first surprise might be that the total cross section is not infinite even though the potential does extend to infinity despite falling rapidly with r. This is because the tails of the potential are not captured by the fluctuating phase, $e^{i\vec{q}\cdot\vec{r}}$, in the Fourier transform. For higher k, the phase oscillates more throughout the integral, hence as $k \to \infty$ the total cross section tends to zero. Another surprise is that the cross section approaches zero as $k \to 0$. But, one must remember that the Born approximation is perturbative and become increasingly invalid as the strength of the potential, V_0, becomes larger than the kinetic energy.

Example 7.2: Born Approximation for the Coulomb Potential
The potential for a charge e scattering off a charge Ze is

$$\mathcal{V}(\vec{r}) = \frac{Ze^2}{r},$$

and the cross section will be the same whether the potential is attractive of repulsive because the matrix element is squared.
Performing the Fourier transform,

$$\begin{aligned}
\tilde{V}(\vec{q}) &= 2\pi \int r^2 dr \int_{-1}^{1} dx e^{iqrx} \frac{Ze^2}{r} \\
&= \frac{4\pi Ze^2}{q} \int dr \sin qr \\
&= -\frac{4\pi Ze^2}{q^2} \cos(qr)|_0^\infty \\
&= \frac{4\pi Ze^2}{q^2} \\
&= \frac{\pi Ze^2}{k^2 \sin^2(\theta/2)},
\end{aligned}$$

The limit at $r \to \infty$ can be realized by adding an exponential damping term to the potential $e^{-\eta r}$, $\eta \to 0^+$, and repeating the integral to see that one gets the same answer with the evaluation at ∞ going to zero.
The expression for the differential cross section, is then,

$$\frac{d\sigma}{d\Omega} = \frac{m^2 Z^2 e^4}{4(\hbar k)^4 \sin^4(\theta/2)}.$$

This answer is identical to the Rutherford cross section in classical mechanics. Note that this cross section is ill-behaved at $\theta \to 0$. It is a non-integrable singulatrity and the total cross section is infinite.

7.3 Structure Functions

Often scattering is performed within a material for the purpose of learning about the structure of the material rather than understanding the potential due to a single scatterer. In that case the Fourier transform of the net potential, i.e. the potential from a large number of scatterers, can be written as,

$$\begin{aligned}\tilde{\mathcal{V}}(\vec{q}) &= \sum_a \int d^3r e^{i\vec{q}\cdot\vec{r}} \mathcal{V}(\vec{r}-\vec{a}) \\ &= \sum_a e^{i\vec{q}\cdot\vec{a}} \int d^3r e^{i\vec{q}\cdot(\vec{r}-\vec{a})} \mathcal{V}(\vec{r}-\vec{a}) \\ &= \tilde{v}(\vec{q}) s(\vec{q}),\end{aligned} \tag{7.16}$$

where

$$s(\vec{q}) \equiv \sum_a e^{i\vec{q}\cdot\vec{a}}, \tag{7.17}$$

and $\tilde{v}(\vec{q})$ is the Fourier transform of the potential from a single scattering center placed at the origin.

This allows us to write the differential cross section,

$$\frac{d\sigma}{d\Omega} = \frac{1}{N}\frac{m^2}{4\pi^2\hbar^4}|\tilde{v}(\vec{q})|^2\,|s(\vec{q})|^2, \tag{7.18}$$

where N is the number of scatterers. The factor $1/N$ enters because the cross sections are related to the rate per scatterer. If the potential for a single scatterer is understood, one may therefore determine $s(\vec{q})$.

The differential cross section is related to the probability that two scattering centers are separated by a distance $\vec{a}$,

$$\begin{aligned}|s(\vec{q})|^2 &= \sum_{a,a'} e^{i\vec{q}\cdot(\vec{a}-\vec{a}')} \\ &= N\sum_{\delta\vec{a}} e^{i\vec{q}\cdot\delta\vec{a}} \\ &= N\tilde{S}(\vec{q}), \\ \frac{d\sigma}{d\Omega} &= \frac{m^2}{4\pi^2\hbar^4}|\tilde{v}(\vec{q})|^2\tilde{S}(\vec{q}).\end{aligned} \tag{7.19}$$

Here, $\tilde{S}(\vec{q})$ is the Fourier transform of the *structure function*,

$$\begin{aligned}\tilde{S}(\vec{q}) &= \frac{d\sigma/d\Omega}{d\sigma/d\Omega|_{\text{isolated}}} \\ &= \sum_{\delta\vec{a}} e^{i\vec{q}\cdot\delta\vec{a}}.\end{aligned} \tag{7.20}$$

This is the ratio of the measured cross section of target particles in matter to the cross section one would measure if the target particles were measured in isolation, i.e. if there was no interference contributions from different target particles. The sum over $\boldsymbol{\delta\vec{a}}$ include $\boldsymbol{\delta\vec{a}} = \mathbf{0}$, i.e. the particle with itself. For large momentum transfers the arguments of the phase factors become large and the terms, $\boldsymbol{e^{i\vec{q}\cdot\delta\vec{a}}}$, become random. Only the $\boldsymbol{\delta\vec{a}}$ term survives and the $\tilde{\boldsymbol{S}}(\vec{\boldsymbol{q}})$ can be set to unity.

One can change the sum over relative positions to an integral where the probability of finding a second scattering center at relative position $\boldsymbol{\delta\vec{a}}$ is $\boldsymbol{S(\delta\vec{a})d^3a}$,

$$\sum_{\boldsymbol{\delta\vec{a}}} \boldsymbol{e^{i\vec{q}\cdot\delta\vec{a}}} \rightarrow \int \boldsymbol{S(\delta\vec{a})d^3a\, e^{i\vec{q}\cdot\delta\vec{a}}}, \tag{7.21}$$

$$\tilde{\boldsymbol{S}}(\vec{\boldsymbol{q}}) = \int \boldsymbol{d^3r\, e^{i\vec{q}\cdot\vec{r}} S(\vec{r})}.$$

The function $\boldsymbol{S}(\vec{\boldsymbol{r}})$ is the structure function in coordinate space. Even in a perfect crystal particles are oscillating around their lattice sites. Thus, a differential density, $\boldsymbol{S}(\boldsymbol{\delta\vec{a}})$, is more realistic than considering a sum over discrete positions, $\boldsymbol{\delta\vec{a}}$.

For a cubic crystal, $\boldsymbol{S}(\vec{\boldsymbol{r}})$ has spikes when $\vec{\boldsymbol{r}} = \boldsymbol{n_x a\hat{x} + n_y a\hat{y} + n_z a\hat{z}}$ where $\boldsymbol{a}$ is the separation between lattice sites. Then, $\tilde{\boldsymbol{S}}(\vec{\boldsymbol{q}})$ has spikes for values of $\vec{\boldsymbol{q}}$ which correspond to in-phase contributions from different sites. For a perfect crystal, these spikes can be infinitely sharp because for certain values of $\vec{\boldsymbol{q}}$ because the contributions can stay in phase as $\vec{\boldsymbol{r}} \rightarrow \boldsymbol{\infty}$. This coherence is known as Bragg's law, https://en.wikipedia.org/wiki/Bragg%27s_law. Whereas, if the material is a liquid, the spikes are greatly muted because there is no long range order. For a liquid, $\boldsymbol{S}(\vec{\boldsymbol{r}})$ typically has a hole near $\vec{\boldsymbol{r}} = \mathbf{0}$ and perhaps a few wiggles before being flat for large $\vec{\boldsymbol{r}}$. The momentum transfer in nuclear or particle experiments is typically high compared to the inverse inter-atomic distance between scatterers. Thus, one often neglects the structure of the target in those experiments. Structure comes into play for X-ray scattering or for low-energy neutron scattering because the characteristic momentum transfers are smaller. However, even with a high energy beam, structure can come into play in the limit that the momentum transfer is sufficiently small, i.e. the scattering is nearly forward.

The considerations used to derive Eq.s (7.19) included an implicit assumption, that the scattering off one target particle at $\vec{\boldsymbol{a}}$ cannot be distinguished from the scattering off a different particle at $\vec{\boldsymbol{a}}'$. If the two target particles are in a lattice, and if the scattering is elastic, the lattice is unchanged by the scattering and the assumption is justified. For higher momentum transfers most of the scattering becomes inelastic, and measuring the structure function becomes more difficult.

Example 7.3: Two Scattering Centers
Consider two scattering centers, one placed at $\vec{\boldsymbol{r}} = \mathbf{0}$ and the second at $\vec{\boldsymbol{r}} = \boldsymbol{a\hat{z}}$, where the beam is in the $\boldsymbol{z}-$direction. Assuming the two scatterers contribute identically, and that the scatterings can be treated perturbatively, at what scattering angles does the differential cross section disappear. Assume the incoming wave number is $\boldsymbol{k\hat{z}}$.

Solution: The overall scattering matrix element has the form

$$\begin{aligned}\alpha(\vec{q}) &= \tilde{v}(\vec{q})\tilde{s}(\vec{q}),\\ \tilde{s}(\vec{q}) &= \sum_a e^{i\vec{q}\cdot\vec{a}},\\ \vec{q} &= |k|((1-\cos\theta)\hat{z} - \sin\theta\cos\phi\hat{x} - \sin\theta\sin\phi\hat{y}).\end{aligned}$$

Here, $\tilde{v}$ is the matrix element for a single scatterer. The cross sections will disappear for directions where $\tilde{s}(\vec{q}) = 0$. This requires

$$\begin{aligned}1 + e^{i\vec{q}\cdot a\hat{z}} &= 0,\\ ka(1-\cos\theta) &= (2n+1)\pi,\end{aligned}$$

where n is a positive integer. This gives an angle for each n,

$$\begin{aligned}\cos\theta_n &= 1 - \frac{(2n+1)\pi}{ka},\\ \theta_n &= \cos^{-1}\left(1 - \frac{(2n+1)\pi}{ka}\right).\end{aligned}$$

For small ka there are no angles at which the cross section vanishes because the arguments of the inverse cosine function must be between ± 1. As ka increases, the number of solutions increases.

7.4 Form Factors

Even if there is only one scattering center, it can be considered as having contributions from many individual points within the single target. If the positions of the points are described by a continuous distribution $\rho(\delta\vec{a})$, one can set $N = 1$ above and replace the sums in Eq. (7.19) with integrals,

$$|F(\vec{q})|^2 = \int d\vec{a}d\vec{a}'\rho(\vec{a})\rho(\vec{a}')e^{i\vec{q}\cdot(\vec{a}-\vec{a}')}. \tag{7.22}$$

Here, ρ is the density of scattering centers such that $\int d\vec{a}\,\rho(\vec{a}) = 1$. With this definition, $|F(\vec{q}=0)|^2 = 1$. The differential cross section is again a product of the cross section one would have if the scattering centers were point-like and the factor $|F(\vec{q}=0)|^2 = 1$, which in the context of scattering off individual object is known as a form factor, $F(\vec{q}=0)$,

$$\frac{d\sigma}{d\Omega} = \left.\frac{d\sigma}{d\Omega}\right|_{\text{point like}} |f(\vec{q})|^2. \tag{7.23}$$

For Coulomb scattering of electrons, the point-like differential cross section is simply the Rutherford form. Then, by measuring the true cross section one determines the form factor, $s(\vec{q})$. By Fourier transforming, on then extracts information about ρ, which describes how charge is distributed about the scattering object.

Example 7.4: Form Factor for a Gaussian Distribution of Charge
Assume the charge density of a proton is distributed as a Gaussian,

$$\rho_q(\vec{r}) = \frac{e}{(2\pi a^2)^{3/2}} e^{-r^2/2a^2}.$$

what is the squared form factor?
Solution: For the form factor define $\rho(\vec{r}) = \rho_q/e$, because it represents the probability the charge is found at $\vec{r}$, i.e., the point-like cross section already has the total charge accounted for, just not its distribution. This gives

$$\begin{aligned}\tilde{F}(\vec{q}) &= \int d^3r \, \frac{1}{(2\pi a^2)^{3/2}} e^{-r^2/2a^2} e^{i\vec{q}\cdot\vec{r}} \\ &= e^{-q^2a^2/2}, \\ |F(\vec{q})|^2 &= e^{-q^2a^2}.\end{aligned}$$

7.5 Higher Order Expansions

Calculating decays and scattering cross sections are the main instances where one invokes Fermi's golden rule. Fermi's golden rule is based on a first-order perturbative correction to the evolution matrix. Higher-order perturbative corrections can be included by replacing the potential V with a more sophisticated operator.

Writing the transition matrix element to second order in time-dependent perturbation theory,

$$\begin{aligned}\langle n|U_I(t,-\infty)|i\rangle|_{n\neq i} &= -\frac{i}{\hbar} V_{ni} \int_{-\infty}^{t} dt' e^{i(\epsilon_n-\epsilon_i)t'/\hbar+\eta t'} \\ &+ \left(\frac{-i}{\hbar}\right)^2 V_{nm}V_{mi} \int_{-\infty}^{t} dt' \int_{-\infty}^{t'} dt'' e^{i(\epsilon_n-\epsilon_m)t'+\eta t'} e^{i(\epsilon_m-\epsilon_i)t''/\hbar+\eta t''} \\ &= -\frac{i}{\hbar} \int_{-\infty}^{t} dt' \, e^{i(\epsilon_n-\epsilon_i)/\hbar+\eta t'} \\ &\cdot \left(V_{ni} + \frac{-i}{\hbar} e^{\eta t'} V_{nm}V_{mi} \int_{-\infty}^{t'} dt'' e^{i(\epsilon_i-\epsilon_m)(t'-t'')/\hbar+\eta(t''-t')} \right).\end{aligned} \tag{7.24}$$

By inspection, one can see that by making the substitution,

$$\begin{aligned}V_{ni} &\to V_{ni} - \frac{i}{\hbar} V_{nm}V_{mi} \int_{-\infty}^{t'} dt'' e^{i(\epsilon_i-\epsilon_m)(t''-t')/\hbar-\eta(t'-t'')} \\ &= V_{ni} - \left.\frac{V_{nm}V_{mi}}{\epsilon_i-\epsilon_m-i\hbar\eta}\right|_{\eta\to 0},\end{aligned} \tag{7.25}$$

one sees that the second order solution looks like the first order solution, but with replaced with the more complicated form, which is known as the T-matrix, T. Sometimes, T refers to some

given order in perturbation theory, but in some contexts might refer to all orders. To all orders the $\boldsymbol{T}$ becomes, extending the derivation for second order in Eq.s (7.24) and (7.25),

$$\begin{aligned} \boldsymbol{T}_{ni} &= \boldsymbol{V}_{ni} + \sum_m \frac{\boldsymbol{V}_{nm}\boldsymbol{V}_{ni}}{(\boldsymbol{E}_m - \boldsymbol{E}_i + i\hbar\boldsymbol{\eta})} \\ &\quad + \sum_{mm'} \frac{\boldsymbol{V}_{nm}\boldsymbol{V}_{mm'}\boldsymbol{V}_{m'i}}{(\boldsymbol{E}_m - \boldsymbol{E}_i + i\hbar\boldsymbol{\eta})(\boldsymbol{E}_{m'} - \boldsymbol{E}_i + i\hbar\boldsymbol{\eta})} + \cdots \end{aligned} \tag{7.26}$$

7.6 Propagators (Time-Ordered Evolution Operators)

Propagators, also known as a Green's functions, are the building blocks of perturbative scattering approaches. A propagator is nothing more than the evolution operator with a step function tacked on,

$$\begin{aligned} \boldsymbol{G}(\boldsymbol{t}) &\equiv \boldsymbol{e}^{-i\boldsymbol{H}\boldsymbol{t}/\hbar}\boldsymbol{\Theta}(\boldsymbol{t}), \\ \boldsymbol{g}(\boldsymbol{t}) &\equiv \boldsymbol{e}^{-i\boldsymbol{H}_0\boldsymbol{t}/\hbar}\boldsymbol{\Theta}(\boldsymbol{t}). \end{aligned} \tag{7.27}$$

The step function is added to account for the time-orderings involved in the expression for the evolution operator in the interaction representation. Remembering that the evolution operator in the interaction representation is

$$\begin{aligned} \boldsymbol{U}_{\boldsymbol{I}}(\boldsymbol{t}-\boldsymbol{t}_0) &= \boldsymbol{e}^{i\boldsymbol{H}_0(\boldsymbol{t}-\boldsymbol{t}_0)/\hbar}\boldsymbol{e}^{-i\boldsymbol{H}(\boldsymbol{t}-\boldsymbol{t}_0)/\hbar} \\ &= \boldsymbol{1} + \left(\frac{-\boldsymbol{i}}{\hbar}\right)\int_{\boldsymbol{t}_0}^{\boldsymbol{t}} \boldsymbol{dt_1}\boldsymbol{e}^{i\boldsymbol{H}_0\boldsymbol{t}_1/\hbar}\boldsymbol{V}\boldsymbol{e}^{-i\boldsymbol{H}_0\boldsymbol{t}_1/\hbar} \\ &\quad + \left(\frac{-\boldsymbol{i}}{\hbar}\right)^{\boldsymbol{2}}\int_{\boldsymbol{0}}^{\boldsymbol{t}} \boldsymbol{dt_1}\boldsymbol{e}^{i\boldsymbol{H}_0\boldsymbol{t}_1}\boldsymbol{V}\boldsymbol{e}^{-i\boldsymbol{H}_0\boldsymbol{t}_1}\int_{\boldsymbol{0}}^{\boldsymbol{t}_1} \boldsymbol{dt_2}\boldsymbol{e}^{i\boldsymbol{H}_0\boldsymbol{t}_2/\hbar}\boldsymbol{V}\boldsymbol{e}^{-i\boldsymbol{H}_0\boldsymbol{t}_2/\hbar} + \cdots, \end{aligned} \tag{7.28}$$

one can write the propagator as

$$\begin{aligned} \boldsymbol{G}(\boldsymbol{t}-\boldsymbol{t}_0) &= \boldsymbol{e}^{-i\boldsymbol{H}_0(\boldsymbol{t}-\boldsymbol{t}_0)/\hbar}\boldsymbol{U}_{\boldsymbol{I}}(\boldsymbol{t})\boldsymbol{\Theta}(\boldsymbol{t}-\boldsymbol{t}_0) \\ &= \boldsymbol{e}^{-i\boldsymbol{H}_0(\boldsymbol{t}-\boldsymbol{t}_0)/\hbar}\boldsymbol{\Theta}(\boldsymbol{t}-\boldsymbol{t}_0) \\ &\quad + \left(\frac{-\boldsymbol{i}}{\hbar}\right)\int_{-\infty}^{\infty} \boldsymbol{dt_1}\, \boldsymbol{e}^{-i\boldsymbol{H}_0(\boldsymbol{t}-\boldsymbol{t}_1)/\hbar}\boldsymbol{\Theta}(\boldsymbol{t}-\boldsymbol{t}_1)\boldsymbol{V}\boldsymbol{e}^{i\boldsymbol{H}_0(\boldsymbol{t}_1-\boldsymbol{t}_0)/\hbar}\boldsymbol{\Theta}(\boldsymbol{t}_1-\boldsymbol{t}_0) \\ &\quad + \left(\frac{-\boldsymbol{i}}{\hbar}\right)^{\boldsymbol{2}}\int_{-\infty}^{\infty} \boldsymbol{dt_1 dt_2}\, \boldsymbol{e}^{-i\boldsymbol{H}_0(\boldsymbol{t}-\boldsymbol{t}_1)/\hbar}\boldsymbol{\Theta}(\boldsymbol{t}-\boldsymbol{t}_1)\boldsymbol{V}\boldsymbol{e}^{i\boldsymbol{H}_0(\boldsymbol{t}_1-\boldsymbol{t}_2)/\hbar}\boldsymbol{\Theta}(\boldsymbol{t}_1-\boldsymbol{t}_2) \\ &\qquad\qquad \cdot \boldsymbol{V}\boldsymbol{e}^{i\boldsymbol{H}_0(\boldsymbol{t}_2-\boldsymbol{t}_0)/\hbar}\boldsymbol{\Theta}(\boldsymbol{t}_2-\boldsymbol{t}_0) \\ &\quad + \cdots \\ &= \boldsymbol{g}(\boldsymbol{t}-\boldsymbol{t}_0) + \left(\frac{-\boldsymbol{i}}{\hbar}\right)\int_{-\infty}^{\infty} \boldsymbol{dt_1}\, \boldsymbol{g}(\boldsymbol{t}-\boldsymbol{t}_1)\boldsymbol{V}\boldsymbol{g}(\boldsymbol{t}-\boldsymbol{t}_0) \\ &\quad + \left(\frac{-\boldsymbol{i}}{\hbar}\right)^{\boldsymbol{2}}\int_{-\infty}^{\infty} \boldsymbol{dt_1 dt_2}\, \boldsymbol{g}(\boldsymbol{t}-\boldsymbol{t}_1)\boldsymbol{V}\boldsymbol{g}(\boldsymbol{t}_1-\boldsymbol{t}_2)\boldsymbol{V}\boldsymbol{g}(\boldsymbol{t}_2-\boldsymbol{t}_0) \\ &\quad + \left(\frac{-\boldsymbol{i}}{\hbar}\right)^{\boldsymbol{3}}\int_{-\infty}^{\infty} \boldsymbol{dt_1 dt_2 dt_3}\, \boldsymbol{g}(\boldsymbol{t}-\boldsymbol{t}_1)\boldsymbol{V}\boldsymbol{g}(\boldsymbol{t}_1-\boldsymbol{t}_2)\boldsymbol{V}\boldsymbol{g}(\boldsymbol{t}_2-\boldsymbol{t}_3)\boldsymbol{V}\boldsymbol{g}(\boldsymbol{t}_3-\boldsymbol{t}_0) + \cdots \end{aligned} \tag{7.29}$$

This expression has a nearly identical form to that for the evolution operator, the only benefit being that the incorporation of the Θ functions allows one to eliminate the limits on the integrals over time by absorbing the limits into the definitions of g. The real benefits to this formalism comes when one considers the Fourier transform of G in frequency.

7.7 The Fourier Transform of the Propagator, $\tilde{G}(\omega)$

Let us first consider the Fourier transform of g,

$$\tilde{g}(\omega) \equiv \int dt\, e^{i\omega t} g(t), \tag{7.30}$$
$$g(t) = \int \frac{d\omega}{2\pi}\, e^{-i\omega t} \tilde{g}(\omega).$$

Assuming the states $|n\rangle$ and $|m\rangle$ are eigenstates of H_0,

$$\begin{aligned}\tilde{g}_{mn}(\omega) &= \delta_{nm} \int dt\, e^{i(\omega - i\epsilon_n/\hbar)t - \eta t}\Theta(t) \\ &= \frac{i\delta_{nm}}{\omega - \epsilon_n/\hbar + i\eta},\ \eta \to 0.\end{aligned}$$

The infinitesimal $\eta \to 0^+$ represents an extremely slow exponential decay at large time differences. By using Cauchy's theorem, one can integrate $\tilde{g}(\omega)e^{-i\omega t}/(2\pi)$ over all ω to see that the propagator $g(t)$ is recovered. In fact, if one flipped the sign of η to 0^-, the result would look the same except with $\Theta(-t)$.

The pole in $\tilde{g}(\omega)$ tells us at what frequency the particle propagates. In this case one sees that the pole (where the propagator blows up) is at $\omega = \epsilon_n/\hbar - i\eta$. If η were replaced with a finite value $\Gamma/2$, the square of the propagator would behave as $e^{-\Gamma t/\hbar}$, and $\Gamma/\hbar$ would be associated with the exponential decay constant.

The advantage to using the operators frequency, rather than as integrals over time, is that the perturbative expansion becomes simpler,

$$\begin{aligned}\tilde{G}(\omega) &= \int dt e^{i\omega(t-t_0)} g(t-t_0) + \left(\frac{-i}{\hbar}\right)\int dt e^{i\omega(t-t_0)} \int dt' g(t-t') V G(t'-t_0) \\ &= \tilde{g}(\omega) + \left(\frac{-i}{\hbar}\right)\int \frac{dt dt' d\omega' d\omega''}{(2\pi)^2} \tilde{g}(\omega') V \tilde{G}(\omega'') e^{i\omega(t-t_0) - i\omega'(t-t') - i\omega''(t'-t_0)} \\ &= \tilde{g}(\omega) + \left(\frac{-i}{\hbar}\right)\tilde{g}(\omega) V \tilde{G}(\omega) \\ &= \tilde{g}(\omega) + \tilde{g}(\omega)\sum_{n>0}\left(\frac{-i}{\hbar}\right)^n (V\tilde{g}(\omega))^n .\end{aligned} \tag{7.31}$$

This is the Dyson series in frequency space, which unlike the one for time, does not involve integrals. This simplification was made possible by incorporating the term $\Theta(t-t')$ into the definition of the Green's function so that the limits of all the integrations over time in Eq. (7.29) could run over all times. Although one typically chooses a basis so that $g(t)$ and $\tilde{g}(\omega)$ are diagonal, V is not usually diagonalized in that basis.

7.8 Propagating to the Same State

The propagator, $\boldsymbol{G_{mn}(\omega)}$, carries information of how one evolves into the state $\boldsymbol{n}$ from the state $\boldsymbol{m}$. The diagonal part of that matrix, $\boldsymbol{G_{mm}(\omega)}$, describes how one remains in, or returns to, the same state, even if the path it involves leaving the state $\boldsymbol{m}$ during intermediate times. By itself, $\boldsymbol{G_{mm}(\omega)}$ provides information on how a particle's energy is modified by the interaction, and how it decays. These considerations form the basis for numerous phenomenological applications of scattering and for numerous descriptions of how particle are modified in medium.

If $\boldsymbol{V}$ is diagonal in the same basis as $\boldsymbol{H_0}$, one can forego the matrix notation and consider the states one-at-a-time. One can then write the expression above for $\tilde{\boldsymbol{G}}$ in a recursive form,

$$\begin{aligned}\tilde{G}_{mm}(\omega) &= \tilde{g}_{mm}(\omega) + \frac{-i}{\hbar}\tilde{g}_{mm}(\omega)V_{mm}\tilde{G}_{mm}(\omega) \\ &= \frac{\tilde{g}_{mm}(\omega)}{1 + (i/\hbar)\tilde{g}_{mm}(\omega)V_{mm}} \\ &= \frac{i}{\omega - \epsilon_m/\hbar - V_{mm}/\hbar + i\eta}.\end{aligned} \tag{7.32}$$

Thus, the only difference between the full propagator $\tilde{\boldsymbol{G}}$ and the original propagator $\tilde{\boldsymbol{g}}$ is that the energy of the pole is shifted by $\boldsymbol{V}$, exactly as one would expect from knowing the eigenvalues of $\boldsymbol{H = H_0 + V}$.

One can not calculate $\boldsymbol{G}$ so easily when $\boldsymbol{V}$ has off-diagonal elements. However, by inspection of the expansion of $\tilde{\boldsymbol{G}}$ in the Dyson expansion, Eq. (7.31), one can separate out the contributions from off-diagonal elements in expansion by defining the matrix $\boldsymbol{\mathcal{T}}$.

$$\begin{aligned}\mathcal{T}_{km}(\omega) \equiv &\sum_{i'\neq m} V_{km} + \left(\frac{-i}{\hbar}\right) V_{ki'}\tilde{g}_{i'i'}(\omega)V_{i'm} \\ &+ \sum_{i'\neq m, j'\neq m}\left(\frac{-i}{\hbar}\right)^2 V_{ki'}\tilde{g}_{i'i'}(\omega)V_{i'j'}\tilde{g}_{j'j'}(\omega)V_{j'm} + \cdots .\end{aligned} \tag{7.33}$$

Thus, $\boldsymbol{\mathcal{T}}$ absorbs all the off-diagonal terms in the expansion of $\boldsymbol{G}$,

$$\begin{aligned}\tilde{G}_{mm}(\omega) &= \tilde{g}_{mm}(\omega) + \left(\frac{-i}{\hbar}\right)\tilde{g}_{mm}(\omega)\mathcal{T}_{mm}(\omega)\tilde{g}_{mm}(\omega) \\ &\quad + \left(\frac{-i}{\hbar}\right)^2 \tilde{g}_{mm}(\omega)\mathcal{T}_{mm}(\omega)\tilde{g}_{mm}(\omega)\mathcal{T}_{mm}(\omega)\tilde{g}_{mm} + \cdots \\ &= \frac{i}{\omega - \epsilon_m/\hbar - \mathcal{T}_{mm}/\hbar + i\eta}.\end{aligned} \tag{7.34}$$

This propagator sums over all paths from some initial state $\boldsymbol{m}$ back into the same final state. All the forays into states other than $\boldsymbol{m}$ are absorbed into $\boldsymbol{\mathcal{T}}$.

The first order correction to $\boldsymbol{\mathcal{T}}$ is simply $\boldsymbol{\mathcal{T}_{mm} = V_{mm}}$. The second order correction in the

potential, separated into real and imaginary parts, is

$$\begin{aligned}\mathcal{T}_{mm}(\omega) &= \sum_{i'\neq m} V_{mi'}\frac{1}{\hbar\omega-\epsilon_{i'}+i\hbar\eta}V_{i'm} \\ &= \sum_{i'\neq m}\left\{\frac{V_{mi'}V_{i'm}}{\hbar\omega-\epsilon_{i'}}+\pi i\delta(\hbar\omega-\epsilon_{i'})V_{mi'}V_{i'm}\right\}.\end{aligned} \tag{7.35}$$

Here we have used

$$\frac{1}{\omega-E/\hbar+i\eta}=\frac{\mathcal{P}}{\omega-E/\hbar}+\pi i\delta(\omega-E/\hbar)\text{ as }\eta\to 0, \tag{7.36}$$

The symbol $\mathcal{P}$ simply points to the fact that the denominator is missing the $i\eta$, and that when integrating across the pole one confines the integral to being away from the singularty by some infinitesimal amount $\pm\epsilon$. Because the part of the integrand that explodes as $1/(\omega-E/\hbar)$, which switches sign as it passes the singularity, the singular part of the integral will not contribute as $\epsilon\to 0$. One can now view the expression for $\tilde{G}_{mm}(\omega)$ for $\omega=\epsilon_m$ to see how the propagator is affected to second order,

$$\begin{aligned}\tilde{G}_{mm}(\omega) &= \frac{i\hbar}{\hbar\omega-\epsilon_m-\Delta\epsilon_m+i\hbar\Gamma_m/2}, \\ \Delta\epsilon_m &= V_{mm}-\sum_{i'\neq m}\frac{V_{mi'}V_{i'm}}{\epsilon_{i'}-\epsilon_m}, \\ \Gamma_m &= \frac{2\pi}{\hbar}\sum_{i'\neq m}V_{mi'}V_{i'm}\delta(\epsilon_m-\epsilon_{i'}).\end{aligned} \tag{7.37}$$

Thus, a pole of the propagator is adjusted by the interaction in such a way that the real part of the pole moves by an amount consistent with stationary state perturbation theory, while the imaginary part is consistent with Fermi's golden rule for calculating the rate at which one of the states decays into the other channel.

7.9 Self-Energy of a Particle in Medium

If the propagator in Eq. (7.37) refers to momentum states, the propagation into the same momentum state k is

$$\begin{aligned}\tilde{G}_{kk}(\omega) &= \frac{i\hbar}{\hbar\omega-\epsilon_k-\Delta\epsilon_k+i\hbar\Gamma_m/2}, \\ \Delta\epsilon_k &= V_{kk}-\sum_{k'\neq k}\frac{V_{kk'}V_{k'k}}{\epsilon_{k'}-\epsilon_k}, \\ \Gamma_k &= \frac{2\pi}{\hbar}\sum_{k'\neq k}V_{kk'}V_{k'k}\delta(\epsilon_k-\epsilon_{k'}).\end{aligned} \tag{7.38}$$

The matrix element V_{kk} has the form,

$$\begin{aligned} V_{kk} &= \langle k|V|k\rangle = \frac{1}{W}\int d^3r\, e^{i\vec{k}\cdot\vec{r}}\mathcal{V}(\vec{r})e^{-i\vec{k}\cdot\vec{r}} \\ &= \frac{1}{W}\int d^3r\, \mathcal{V}(\vec{r}). \end{aligned} \tag{7.39}$$

Here, the volume of the plane wave, W, usually goes to infinity. This kills the contributions to both $\Delta\epsilon_k$ and Γ_k, and the position of the propagators pole does not change in a scattering experiment. In fact, if one calculated higher-order corrections to $\Delta\epsilon$, they would all scale as $1/W$.

However, if the particle is traversing a medium, where the density of such scatters in n_s, the volume per scatterer becomes $1/n_s$, and the factor $1/W$ becomes n_s. In this case the propagator's pole, essentially the energy of the particle, is altered a finite amount,

$$\Delta\epsilon_k = n_s\int d^3r\mathcal{V}(\vec{r}), \tag{7.40}$$

to first order in perturbation theory. After changing the sum over intermediate states k' to an integral,

$$\sum_{k'} = \frac{W}{(2\pi)^3}\int k'^2 dk d\Omega', \tag{7.41}$$

the decay rate in Eq. (7.38), Γ_k, becomes

$$\begin{aligned} \Gamma_k &= n_s\frac{2\pi}{\hbar(2\pi)^3}\int k'^2 dk' d\Omega'\,\delta(\epsilon_k-\epsilon_{k'})\left|\int d^3r\,\mathcal{V}(\vec{r})e^{i(\vec{k}-\vec{k}')\cdot\vec{r}}\right|^2 \\ &= n_s\frac{mk}{4\pi^2\hbar^3}\int d\Omega'\left|\int d^3r\,\mathcal{V}(\vec{r})e^{i(\vec{k}-\vec{k}')\cdot\vec{r}}\right|^2. \end{aligned} \tag{7.42}$$

Using Eq. (7.13), and replacing $v = \hbar k/m$, this becomes

$$\Gamma_k = n_s\sigma v, \tag{7.43}$$

which is the scattering rate in a medium of density n_s.

The correction to the particle's energy, $\Delta\epsilon_k$, is known as the real part of the particles' *self energy*, and $\hbar\Gamma_k/2$ is the imaginary part. The change in the pole of the propagator due to the interaction represents the change in the wave function of a particle of momentum $\vec{p}$,

$$e^{-i\epsilon^{(0)}(p)t/\hbar+i\vec{p}\cdot\vec{r}/\hbar}\to e^{-i(\epsilon^{(0)}(p)+\Delta\epsilon(p))t/\hbar+i\vec{p}\cdot\vec{r}-\Gamma(p)t/2}. \tag{7.44}$$

Squaring the wave function results in $e^{-\Gamma(p)t}$, illustrating the decaying nature of the wave in the medium. In addition to the energy changing the velocity also changes in the medium. The velocity of a wave is given by the expression,

$$v_i = \frac{d(\epsilon^{(0)}(p)+\Delta\epsilon(p)}{dp_i}. \tag{7.45}$$

For example, if photons of energy ϵ are traversing a medium where they might excite molecules into some excited state E^*, the correction $\Delta\epsilon$ might be positive or negative depending on whether the excited state is above or below ϵ. This switch of sign results from the denominator of the expression for $\Delta\epsilon$ changing sign if the energy of the intermediate state passes ϵ. The group velocity follows suit, and can be significantly altered for $\epsilon \approx E^*$. Of course, when $\epsilon \approx E^*$, the cross section, and thus the decay rate, is also large. This won't be pursued further in this course, but it is clear that the self energy and its momentum dependence play a critical role in understanding how a particle's properties are altered in medium.

7.10 Resonant Scattering

Here we consider the problem where an initial state k scatters to a final state k' through a resonant channel R. For simplicity, we consider the matrix element of the momentum state k with the resonant state R to be

$$\langle R|V|k\rangle = \frac{\alpha(k)}{\sqrt{W}}, \tag{7.46}$$

We will assume that one has performed all the necessary integrations to find the matrix element, and that we call it α aside from the $1/\sqrt{W}$ due to the normalization of the plane wave in the volume W. To simplify our discussion we will assume that α has no dependence with respect to the direction of k.

One could calculate the rate of decay of the resonance using Fermi's golden rule and obtain

$$\begin{aligned}\Gamma_R &= \frac{2\pi}{\hbar}\frac{|\alpha|^2}{W}\rho_k(\epsilon_R), \\ &= \frac{k^2}{\pi v_k \hbar^2}|\alpha|^2\end{aligned} \tag{7.47}$$

where $\rho_k(\epsilon_R)$ is the density of states of the outgoing particle.

An example of resonant scattering could be a photon scattering off an atom. Here k labels the momentum of the photon, while R would refer to a specific excited state of the atom that could be attained due to the interaction with the photon. To simplify our derivation, we will assume that α is independent of k.

It was shown in the previous lecture that the cross section for scattering could be written to second order as

$$\sigma = \frac{2\pi V}{v\hbar}\sum_{k'}|\mathcal{T}_{k'k}|^2\delta(\epsilon_{k'}-\epsilon_k) \tag{7.48}$$

$$\mathcal{T}_{k'k} = V_{k'k} + \frac{-i}{\hbar}\sum_m V_{k'm}\frac{i}{\omega - \epsilon_m/\hbar + i\eta}V_{mk}\bigg|_{\hbar\omega=\epsilon_k}.$$

Using the derivations of the previous section, one could have inserted the full propagator into

the intermediate state, which would then yield

$$\mathcal{T}_{k'k} = V_{k'k} + \frac{-i}{\hbar}\sum_{mm'} V_{k'm}\tilde{G}_{mm'}(\omega)V_{m'k}\Big|_{\hbar\omega=\epsilon_k} . \tag{7.49}$$

By resonant scattering we mean that the only interaction is between the resonant state R and the momentum states. The matrix element then becomes

$$\begin{aligned}\mathcal{T}_{k'k} &= \frac{-i}{\hbar} V_{k'R}\tilde{G}_{RR}(\omega)V_{Rk}\Big|_{\hbar\omega=\epsilon_k} , \\ \tilde{G}_{RR}(\omega) &= \frac{i}{\omega - \epsilon_R/\hbar + i\Gamma_R/2} ,\end{aligned} \tag{7.50}$$

where Γ is the decay rate of the resonance.

The cross section is no longer infinite when $\epsilon_k = \epsilon_R$ due to the factor Γ_R appearing in the denominator. Plugging the expression for $\mathcal{T}$ into the expression for the cross section one obtains

$$\begin{aligned}\sigma &= \frac{2\pi W}{v\hbar}\sum_{k'}\frac{|\alpha|^4}{W^2}\frac{1}{(\epsilon_k-\epsilon_R)^2+\hbar^2\Gamma_R^2/4}\delta(\epsilon_{k'}-\epsilon_k) \\ &= \frac{2\pi W}{v\hbar}\frac{|\alpha|^4}{W^2}\frac{1}{(\epsilon_k-\epsilon_R)^2+\hbar^2\Gamma_R^2/4}\rho_k(E) \\ &= \frac{k^2}{\pi v^2\hbar^2}|\alpha|^4\frac{1}{(\epsilon_k-\epsilon_R)^2+\hbar^2\Gamma_R^2/4} \\ &= \frac{4\pi}{k^2}\frac{(\hbar\Gamma_R/2)^2}{(\epsilon_k-\epsilon_R)^2+(\hbar\Gamma_R/2)^2} .\end{aligned} \tag{7.51}$$

The last line is known as the Breit-Wigner form for scattering through a resonance. Note that the cross section is determined by two numbers, the width of the resonance and the Energy of the resonance.

If the resonance has spin S_R, then the effect is multiplied by the number of degenerate states through which one might scatter. If the incoming particles have spins S_1 and S_2, the effect is correspondingly reduced due to the fact that many of the states of the resonance would not be reached with particular combinations of S_1 and S_2.

The Breit-Wigner form for resonant scattering is then

$$\sigma = \frac{(2S_R+1)}{(2S_1+1)(2S_2+1)}\frac{4\pi}{k^2}\frac{(\hbar\Gamma_R/2)^2}{(\epsilon_k-\epsilon_r)^2+(\hbar\Gamma_R/2)^2} . \tag{7.52}$$

Example 7.5: Resonant Scattering through the ρ Meson
Here, we consider $\pi - \pi$ resonant scattering through the ρ meson. Consider a π_+ and a π_0 which are scalar mesons with masses 140 MeV/c^2 and 135 MeV/c^2. They scatter through the ρ_+ resonance which has spin 1, a mass of 770 MeV/c^2 and a width $\hbar\Gamma = 151$ MeV. Find the cross section at resonance.

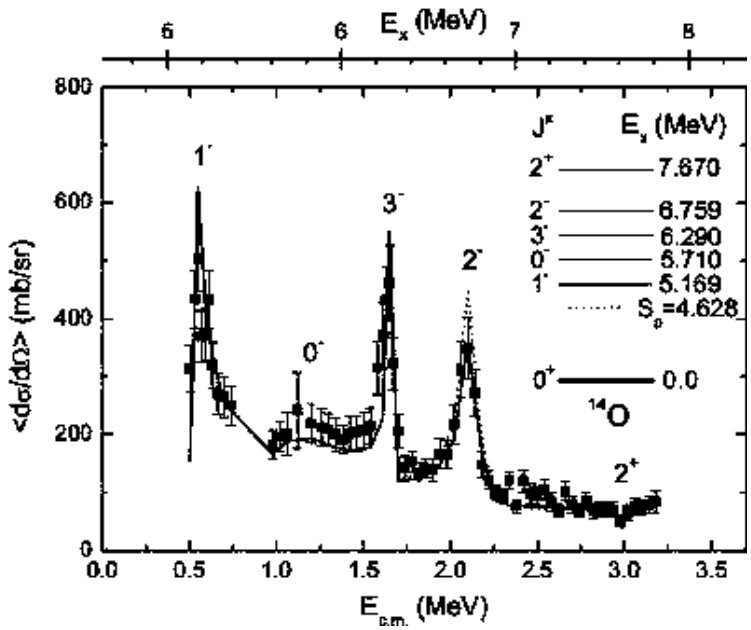


Figure 7.1: An example of resonant scattering of $p+^{13}$C. Each peak corresponds to a state in ^{14}O. Figure from Y. Wang et al., "Study of Elastic Resonance Scattering at Ciae" (2011).

Solution: First, one must find the relative momentum $\vec{p}$ in the center of mass frame. In this frame the particles move with $\vec{p}$ and $-\vec{p}$.

$$m_\rho = \sqrt{m_+^2 + p^2} + \sqrt{m_0^2 + p^2},$$
$$p^2 = \frac{m_\rho^4 + m_+^4 + m_0^4 - 2m_\rho^2 m_+^2 - 2m_\rho^2 m_0^2 - 2m_-^2 m_+^2}{4m_\rho^2}.$$

Here, we have neglected the factors of c. The value of p at resonance is then 360 MeV/c. The cross section at resonance is then

$$\sigma_R = (2S+1)\frac{4\pi}{k_R^2} = \frac{12\pi}{(360/197.326)^2} = 11.3 \text{ fm}^2 = 113 \text{ mb}.$$

Here, $\hbar c = 197.326$MeV·fm and one fm^2 equals 10 millibarns. Note that the maximum cross section depended only on the relative momentum at resonance and not on the width.

7.11 Resonant Scattering, Simplified

Repeating Eq. (7.25), one can extend Fermi's golden rule to second order via the substitution,

$$V_{fi} \to V_{fi} - \frac{i}{\hbar} V_{fR} V_{Ri} \int_{-\infty}^{t'} dt'' e^{i(\epsilon_i - \epsilon_m)(t''-t')/\hbar - \eta(t'-t'')} \tag{7.53}$$
$$= V_{fi} - \frac{V_{fR} V_{Ri}}{\epsilon_i - \epsilon_R - i\hbar\eta}\bigg|_{\eta\to 0}.$$

The second term represents the part of the evolution that began in state i at $t = -\infty$, then switched to state R at time t'', then had the state R decay back to a state f at time t', which due to the delta function in Fermi's golden rule would have the same energy as the state i. One might modify the exponential term in the integral to account for the decay,

$$e^{i(\epsilon_i-\epsilon_m)(t''-t')/\hbar-\eta(t'-t'')} \rightarrow e^{i(\epsilon_i-\epsilon_m)(t''-t')/\hbar-\Gamma(t'-t'')/2}, \tag{7.54}$$

with the decay rate being Γ. The fact that $\Gamma/2$ is being used is motivated by the fact that squaring the amplitude would give the factor $e^{-\Gamma t}$. This ansatz is intuitive, but nonetheless is an assumption and is therefore less rigorous than then derivation of the previous section. Eq. (7.53) now becomes,

$$V_{fi} \rightarrow V_{fi} - \frac{V_{fR}V_{Ri}}{\epsilon_i - \epsilon_R - i\hbar\Gamma/2}. \tag{7.55}$$

If the magnitude of V_{fR} is independent of the direction of the momentum for all states with the same energy, one can replace V_{fR} and V_{Ri} with $|V|$. One can also express the decay rate in terms of V using Fermi's golden rule for the decay,,

$$\Gamma = \frac{2\pi}{\hbar}|V|^2\rho(\epsilon_i), \tag{7.56}$$

where the delta function in Fermi's golden rule has been replaced by the density of states,

$$\begin{aligned}\rho(\epsilon_i) &= \frac{W}{(2\pi)^3}\int 4\pi k^2 dk \delta(\epsilon_k - \epsilon_i) \\ &= W\frac{mk}{2\pi^2\hbar^2},\end{aligned} \tag{7.57}$$

where W is the large volume and k is chosen to satisfy the delta function.

Substituting for V and neglecting the first term for V_{fi},

$$\begin{aligned}V_{fi} &\rightarrow \frac{\hbar\Gamma}{2\pi\rho(\epsilon_f)}\frac{1}{(\epsilon_i - \epsilon_R - i\hbar\Gamma/2)} \\ &= \frac{1}{\pi\rho(\epsilon_f)}\frac{\hbar\Gamma/2}{(\epsilon_i - \epsilon_R - i\hbar\Gamma/2)}.\end{aligned} \tag{7.58}$$

Repeating Eq. (7.9), the cross section expressed in Fermi's golden rule,

$$\sigma = \frac{2\pi W}{v_f\hbar}|V|^2\rho(\epsilon_f). \tag{7.59}$$

Here, $|V|$ will be replaced with the expression in Eq. (7.55). One can now express the cross section in terms of Γ,

$$\begin{aligned}\sigma &= \frac{2W}{v_f\pi\hbar\rho(\epsilon_f)}\frac{(\hbar\Gamma/2)^2}{(\epsilon_i - \epsilon_R)^2 + (\hbar\Gamma/2)^2} \\ &= \frac{4\pi}{k^2}\frac{(\hbar\Gamma/2)^2}{(\epsilon_i - \epsilon_R)^2 + (\hbar\Gamma/2)^2}.\end{aligned} \tag{7.60}$$

The velocity of the outgoing particle was replaced with $mv_f = \hbar k$. This is the same result for the Breit-Wigner cross section in Eq. (7.52) derived more rigorously before.

Example 7.6: Narrow Resonance
Imagine that a proton could decay to a positron and a photon with a long lifetime $\tau = 10^{35}$ years. Rather than waiting for a decay, you build an accelerator that collides photons with electrons with the required center-of-mass energy. To simplify the calculation, you can make the approximation that the electron is massless its energy in the center-of-mass frame, ~ 470 MeV, is much larger than the electron mass.

a) What is the cross section at resonance?

b) If one can tune the beam so that the center-of-mass energy is $m_p \pm 1$ eV, what is the averaged cross section over the energy range?

Solution:
a) The Breit-Wigner cross section, using the degeneracy factors for the proton, photon and electron, is

$$\begin{aligned}
\sigma_R &= \frac{1}{2}\frac{4\pi}{k_R^2} \\
\hbar k_R c &= 470 \text{ MeV}, \\
k_R &= \frac{470}{197.326 \text{MeV fm}} = 2.38 \text{ fm}^{-1}, \\
\sigma_R &= 1.11 \text{ fm} = 11.1 \text{ mb}.
\end{aligned}$$

Collider experiments readily measure cross sections many thousands of times smaller than this.
b) Now, consider the energy-averaged cross section

$$\bar{\sigma} = \frac{1}{2\Delta E}\int_{m_p-\Delta E}^{m_p+\Delta E} dE\, \sigma(E).$$

The width, $\Delta E = 1$ eV, is much larger than the width, $\hbar\Gamma$, of the resonance, so the limits of the integral can be taken to infinity. Further, ΔE is small enough that the prefactor $4\pi/k^2$ can be set to $4\pi k_R^2$. Thus,

$$\begin{aligned}
\bar{\sigma} &= \frac{\sigma_R}{2\Delta E}\int_{-\infty}^{\infty} dE \frac{(\hbar\Gamma/2)^2}{(E-m_p)^2 + (\hbar\Gamma/2)^2} \\
&= \frac{\sigma_R \hbar\Gamma}{4\Delta E}\int_{-\infty}^{\infty} dx\, \frac{1}{x^2+1} \\
&= \frac{\pi\hbar\Gamma}{4\Delta E}\sigma_R.
\end{aligned}$$

Using $\hbar = 6.58 \times 10^{-16}$ eVs, and $\Gamma = 10^{35}$ years $= 3.15^{42}$ s,

$$\hbar\Gamma = \frac{\hbar}{\tau} = 3.17 \times 10^{-59} \text{ eV}.$$

This gives

$$\bar{\sigma} = 2.8 \times 10^{-58} \text{ mb}.$$

This is FAR too small to be measured.

7.12 Exercises

1. Using the Born approximation estimate the differential scattering cross section, $d\sigma/d\Omega$ for particles of mass m scattering off the following potentials.

   (a) $\mathcal{V}(\vec{r}) = V_0\Theta(a-r)$.
   (b) $\mathcal{V}(\vec{r}) = a^3V_0\delta^3(\vec{r})$.
   (c) $\mathcal{V}(\vec{r}) = a^3V_0\left[\delta^3(\vec{r}-a\hat{z}) + \delta^3(\vec{r}+a\hat{z})\right]$.
   (d) $\mathcal{V}(\vec{r}) = a^3V_0\left[\delta^3(\vec{r}-a\hat{z}) - \delta^3(\vec{r}+a\hat{z})\right]$.
   (e) $\mathcal{V}(\vec{r}) = a^3V_0\left[\delta^3(\vec{r}-a\hat{x}) - \delta^3(\vec{r}+a\hat{x})\right]$.
   (f) $V_0e^{-r/a}/r$.

   You can express your answer either in terms of the momentum transfer $\vec{q}$, or in terms of the beam momentum k and the scattering angle θ, ϕ.

2. By taking two derivatives of the form factor at $q = 0$,

$$\frac{\partial}{\partial q_i}\frac{\partial}{\partial q_j}F(\vec{q})\bigg|_{q=0},$$

   one can generate moments of the charge distribution,

$$\langle(r_i-\bar{r}_i)(r_j-\bar{r}_j)\rangle \equiv \int d^3r\,\rho(\vec{r})(r_i-\bar{r}_i)(r_j-\bar{r}_j).$$

   (a) Working in a coordinate system where $\bar{r}_i = 0$, prove the relation above.
   (b) Test your answer by comparing to the result of Example 7.4. First calculate $\langle r_ir_j\rangle$ by integrating to get a weighted average of r_ir_j using $\rho_q(\vec{r})$, then compare to the expression above using derivatives of the form factor.

3. The cross section for scattering a particle with momentum $\hbar k$ off a single target is

$$\frac{d\sigma}{d\Omega} = \alpha, \tag{7.61}$$

   which is independent of the scattering angle. Here, we assume that the cross section is small. Now, two targets are placed a distance a apart, separated along the z axis (the same axis along which the incident beam is directed). At what scattering angles does the differential cross section, $d\sigma/d\Omega$, equal zero? (This is worked out as an example in the notes)

4. The cross section for scattering a particle with momentum $\hbar k$ off a single target is

$$\frac{d\sigma}{d\Omega} = \alpha, \tag{7.62}$$

   which is independent of the scattering angle. Here, we assume that the cross section is small. Now, two targets are placed a distance a apart, separated along the x axis (perpendicular to the axis along which the incident beam is directed). At what scattering angles does the differential cross section, $d\sigma/d\Omega$, equal zero? Specify both the polar angle θ and the azimuthal angle ϕ.

5. Consider a charge Z that is uniformly distributed within a sphere of radius R. A
   (a) Find the squared form factor $F(\vec{q})$.
   (b) For an incoming wave with wave number $k\hat{z}$ what is F as a function of the scattering angle θ? Express F in terms of kR and θ.
   (c) Plot $|F|^2$ from (b) for the cases where $kR = 1, 2$ and 5 as a function of θ.

6. A π^+, which is a spin-zero meson, scatters off a proton through a Δ^{++} resonance(which is comprised of three up quarks). The Δ^{++} is spin 3/2 baryon. The masses of the pion, proton and delta are 139.58 MeV/c^2, 938.28 MeV/c^2 and 1232 MeV/c^2 respectively. The width of the Δ is 120 MeV.
   (a) Using relativistic dispersion relations, $E = \sqrt{p^2c^2 + m^2}$, what is the relative momentum, q, of the pion and proton at resonance? This is one half the difference of the two momenta in the center-of-mass frame. I.e. $\epsilon_\pi(p_R) + \epsilon_p(p_R) = M_\Delta$.
   (b) Estimate the cross section at resonance? Give your answer in millibars. One mb equals 10^{-24} cm and 10 mb= 1 fm^2.

7. Consider a particle of mass m that could be confined to a spherical well,

$$\mathcal{V}(r) = \begin{cases} 0, & r < a \\ V_0, & a < r < 2a \\ 0, & r > 2a \end{cases}$$

   (a) Use the WKB method to estimate the decay rate of a particle of mass m escaping from a spherical trap defined by the potential. Assume the barrier is sufficiently high to approximate the energy of the trapped particle with an infinite well.
   (b) Find an expression to estimate the cross section for a particle scattering off the potential well with an energy near the ground state energy described above. You can give your answer as a function of the incident energy, E, m, V_0, a, and the width Γ.

8. Consider the function

$$\tilde{f}(\omega) = \frac{i}{\omega - E/\hbar + i\Gamma/2}.$$

   Show that the Fourier transform is

$$\begin{aligned} f(t) &= \frac{1}{2\pi}\int d\omega\, e^{-i\omega t}\tilde{f}(\omega) \\ &= e^{-iEt/\hbar - \Gamma t/2}\Theta(t). \end{aligned}$$

   HINT: Use contour integration.

8 Scattering at Lower Energies

Perturbative approaches do not work when the potential, which is treated as a perturbation, is larger than the kinetic term. This then invalidates the approaches of the previous chapter. Fortunately, when the kinetic energy is small, the relevant angular momenta are also small. If the range of the potential is a, the angular momenta that contribute to the scattering are in the range, $L \lesssim pa$, where p is the momentum of the beam particles. Many experiments work in the regime where only a few units of angular momenta need to be considered, and in fact, a large amount of information about potentials is extracted from experiments where one need only consider $\ell = 0$.

Of course, the incoming and outgoing measurements of scattering experiments are plane waves, eigenstates of momentum, not angular momentum. However, plane waves can be decomposed into spherical waves (known as partial waves), i.e. waves that are eigenstates of angular momentum, with quantum numbers $\ell = 0, 1, 2 \cdots$. In this chapter, we will treat scattering by considering a few partial waves. The scattering wave is then considered as the original plane wave, plus the alterations of those partial waves with low ℓ. The alterations to the partial waves due to the potential then provide one with expressions for the differential cross section. If the potential is spherically symmetric, one can consider partial waves of a given ℓ, individually. This effectively reduces the three-dimensional scattering problem to a few one-dimensional problems with fixed ℓ.

8.1 Partial Waves and Phase Shifts

If potentials have spherical symmetry, angular momentum is conserved and one can solve the scattering problem through the consideration of spherical waves rather than plane waves. This effectively reduces the problem to the solution of one-dimensional Schrödinger equations, where each partial wave is characterized by a specific angular momentum ℓ and projection m. Each partial requires a different centrifugal potential as the angular momentum barrier depends on ℓ. Because the energies are low, the relative momenta and the angular momenta are small. For low energy scattering only the lowest few ℓ need to be considered, and when the scattering energy approaches zero, scattering is dominated by the $\ell = 0$ contributions.

First, we consider the kinetic part of the Hamiltonian. The Schrödinger equation can be written in either Cartesian coordinates or spherical coordinates, and the kinetic term is

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right) = -\frac{\hbar^2}{2m}\left(\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\right) \tag{8.1}$$

$$= -\frac{\hbar^2}{2m}\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{L^2}{2mr^2}.$$

If angular momentum is conserved, L^2 is a constant and can then be replaced with its eigenvalue $\hbar^2\ell(\ell+1)$. Then $\hbar^2\ell(\ell+1)/2mr^2$ appears like a repulsive potential energy in Schrödinger's equation, and is thus referred to as a centrifugal potential. If the potential is spherically symmetric, angular momentum is a good quantum number and one can write solutions corresponding

to a specific ℓ and m.

$$H\psi_{\ell,m}(\vec{r}) = E\psi_{\ell,m}(\vec{r}), \tag{8.2}$$

with the wave function being written as a product of a radial part and an angular part,

$$\psi_{\ell,m}(\vec{r}) = Y_{\ell,m}(\theta,\phi)R_\ell(r). \tag{8.3}$$

The radial wave function is a solution of the equation

$$-\frac{\hbar^2}{2m}\frac{1}{r}\frac{\partial^2}{\partial r^2}rR_\ell(r) + \frac{\hbar^2\ell(\ell+1)}{2mr^2}R_\ell(r) + V(r)R_\ell(r) = ER_\ell(r). \tag{8.4}$$

Remember that the radial wave function depends only on ℓ and not m because the centrifugal potential is determined by ℓ only.

If the potential is zero, the solutions $R_\ell(r)$ are spherical Bessel functions $j_\ell(kr)$, first presented in Chapter 4. Plane wave solutions can be expanded in terms of the spherical solutions through,

$$e^{i\vec{k}\cdot\vec{r}} = \sum_\ell (2\ell+1)i^\ell j_\ell(kr)P_\ell(\cos\theta), \tag{8.5}$$

where $\cos\theta \equiv \hat{k}\cdot\hat{r}$. This is known as the partial wave expansion. Note that the expansion is in terms of the Legendre polynomials,

$$P_\ell(\cos\theta) = \sqrt{\frac{4\pi}{2\ell+1}}Y_{\ell,m=0}(\theta,\phi). \tag{8.6}$$

All angular functions can be expressed in terms of $Y_{\ell,m}$s, and because $e^{i\vec{k}\cdot\vec{r}} = e^{ikr\cos\theta}$ does not depend on ϕ, it is not surprising that the expansion contains only $Y_{\ell,m=0}$ terms. The partial wave expansion can be derived by combining the orthogonality relations of the Legendre polynomials and spherical Bessel functions with the Rodriquez formula,

$$P_\ell(x) = \frac{1}{2^\ell \ell!}\frac{d^\ell(x^2-1)^\ell}{dx^\ell}. \tag{8.7}$$

The solutions j_ℓ are normalized so that for large $x = kr$ they behave as,

$$j_\ell(x)|_{x\to\infty} = \frac{(-i)^\ell e^{ix} - (i^\ell)e^{-ix}}{2ix}. \tag{8.8}$$

They are a linear combination of outgoing and incoming waves, each of which is a solution to the Schrödinger equation. The relative phase between the incoming and outging waves is chosen so that the solution goes to zero at $x = 0$. In fact,

$$j_\ell(x) \approx \left.\frac{x^\ell}{(2\ell+1)!!}\right|_{x\approx 0}. \tag{8.9}$$

One can verify the small x expansion by applying Schrödinger's equation with the centrifugal potential. Using the definition, $u_\ell(x) \equiv xR_\ell(x)$, we assume that u_ℓ is an expansion in powers of x, with the lowest power being n. Schródingers's equation becomes

$$-\partial_x^2 u_\ell(x) + \frac{\ell(\ell+1)}{x^2}u_\ell(x) = u_\ell(x), \tag{8.10}$$

$$u_\ell(x) = x^n + ax^{n+1} + bx^{n+2} + \cdots$$

In order for the expansion to not include lower powers in m one must have

$$-\frac{n(n-1)}{x^2}+\frac{\ell(\ell+1)}{x^2}=0, \tag{8.11}$$

where the factor $n(n-1)$ comes from taking two derivatives of x^n. This give $n=\ell+1$ or $n=-\ell$. Because $R_\ell \sim u_\ell/r$, the radial wave functions behave as r^ℓ or $r^{-\ell-1}$.

The second set of solutions, those where u_ℓ behaves as $x^{-\ell}$, do not satisfy the boundary conditions at zero as they begin with $x^{-\ell-1}$. These are known as spherical Neumann functions and have the opposite relative phase between the incoming and outgoing parts, while being quite divergent at the origin.

$$n_\ell(x) \approx \frac{(2\ell-1)!!}{x^{\ell+1}}. \tag{8.12}$$

Examples of a few spherical functions for low ℓ, repeated from Chapter 4, are

$$\begin{aligned} j_0(x)&=\frac{\sin x}{x},\ n_0(x)=-\frac{\cos x}{x} \\ j_1(x)&=\frac{\sin x}{x^2}-\frac{\cos x}{x},\ n_1(x)=-\frac{\cos x}{x^2}-\frac{\sin x}{x} \\ j_2(x)&=\left(\frac{3}{x^3}-\frac{1}{x}\right)\sin x-\frac{3}{x^2}\cos x,\ n_2(x)=-\left(\frac{3}{x^3}-\frac{1}{x}\right)\cos x-\frac{3}{x^2}\sin x \end{aligned} \tag{8.13}$$

Both j_ℓ and n_ℓ are real.

By making combinations, $j_\ell \pm i n_\ell$, one finds solutions that correspond to incoming or outgoing waves. The spherical Hankel functions are defined in terms of j_ℓ and n_ℓ,

$$\begin{aligned} h_\ell(x) &\equiv j_\ell(x)+in_\ell(x) \approx \frac{(-i)^{\ell+1}}{x}e^{ix}|_{x\to\infty} \\ h_\ell^*(x) &\equiv j_\ell(x)-in_\ell(x) \approx \frac{i^{\ell+1}}{x}e^{-ix}|_{x\to\infty} \end{aligned} \tag{8.14}$$

Here h_ℓ and h_ℓ^* behave as outgoing and incoming waves respectively.

When adding a potential of finite range, there still exists solutions which look like $h_\ell(kr)$ or $h_\ell^*(kr)$ for r beyond the range of the potential, but have modified forms at small r. Just as in the case with no potential, one can find a linear combination of the incoming and outgoing solutions which goes to zero at $r=0$. However, the relative phase between the incoming and outgoing phase will be adjusted by a phase $e^{2i\delta_\ell}$ due to the existence of the potential. The modification for r outside the range of the potential can only be a phase factor because the incoming flux and outgoing flux must be the same strength. The large $x=kr$ behavior is then

$$R_\ell(x)|_{x>ka}=\frac{1}{2}\left(e^{2i\delta_\ell}h_\ell(x)+h_\ell^*(x)\right), \tag{8.15}$$

where δ is known as the phase shift. Here a is any distance large enough such that the potential is zero (keep in mind that this is never true for the Coulomb potential). We define the overall

phase of R_ℓ so that the incoming phase has the same phase as does the incoming part of j_ℓ. For the $s-$wave, this becomes

$$R_{\ell=0}(k, r > a) = \frac{e^{i\delta}\sin(kr+\delta)}{kr}. \tag{8.16}$$

If one scatters a plane wave off a potential, one can consider the solution to be the original plane wave, expanded in terms of partial waves, plus the correction due to the interaction,

$$\begin{aligned}\psi_{\vec{k}}(\vec{r}) &= \sum_\ell (2\ell+1) i^\ell R_\ell(k,r) P_\ell(\cos\theta) \\ &= e^{i\vec{k}\cdot\vec{r}} + \sum_\ell (2\ell+1) i^\ell \left(R_\ell(k,r) - j_\ell(kr)\right) P_\ell(\cos\theta),\end{aligned} \tag{8.17}$$

where the choice of phases in the definition of R_ℓ allows the incoming waves to be identical to those of the solution with no potential. Expanding the answer at large r, one obtains

$$\begin{aligned}\psi_{\vec{k}}(\vec{r})|_{r\to\infty} &= e^{i\vec{k}\cdot\vec{r}} + \sum_\ell (2\ell+1)\left(e^{2i\delta_\ell}-1\right)\frac{e^{ikr}}{2ikr} P_\ell(\cos\theta) \\ &= e^{i\vec{k}\cdot\vec{r}} + \sum_\ell (2\ell+1) e^{i\delta_\ell}\sin\delta_\ell \frac{e^{ikr}}{kr} P_\ell(\cos\theta)\end{aligned} \tag{8.18}$$

Only the latter term contributes to scattering as the plane wave continues to travel forward after the wave packet leaves the region of the scatterer. One defines a quantity $f(\Omega)$ as the scattering amplitude,

$$\begin{aligned} f(\Omega) &\equiv \sum_\ell (2\ell+1) e^{i\delta_\ell}\sin\delta_\ell \frac{1}{k} P_\ell(\cos\theta) \\ \psi_{\vec{k}}(\vec{r})|_{R\to\infty} &= e^{i\vec{k}\cdot\vec{r}} + \frac{e^{ikr}}{r} f(\Omega).\end{aligned} \tag{8.19}$$

The scattering amplitude, $f(\Omega)$ is a function of the scattering angle θ and has dimensions of length. If there were no spherical symmetry, $f(\Omega)$ could also depend on the azimuthal angle ϕ. Note that f no longer depends on r. It can also be related to the differential cross section. To see this, we first relate the differential cross section to the flux of particles per solid angle,

$$\frac{v}{V}\frac{d\sigma}{d\Omega} = \frac{dN}{d\Omega dt}. \tag{8.20}$$

The flux per unit area can be found by multiplying the square of the wave function in Eq. (8.19) by the velocity and dividing by the volume,

$$\frac{dN}{r^2 d\Omega dt} = \frac{v}{V}\frac{|f(\Omega)|^2}{r^2}, \tag{8.21}$$

Comparing the two equations above allows one to see that $f(\Omega)$ is directly related to the differential cross section.

$$\frac{d\sigma}{d\Omega} = |f(\Omega)|^2. \tag{8.22}$$

Combined with Eq. (8.19), one can see that the differential cross section at a given energy is determined solely by the phase shifts $\delta_\ell(k)$.

Squaring the scattering amplitude in Eq. (8.19) includes cross terms with different values of ℓ, but the contribution from the cross terms disappears after integrating over $d\Omega$ to obtain the total cross section. The cancellation is due to the orthogonality of the Legendre Polynomials. The total cross section is then

$$\sigma = \frac{4\pi}{k^2}\sum_\ell (2\ell+1)\sin^2\delta_\ell. \tag{8.23}$$

For resonances of a given ℓ the peak of the resonance is for where $\sin^2(\delta_\ell(k))$ is near unity. The value of σ at that point, $4\pi/k^2$, then matches the same value for the Breit-Wigner resonance in Eq. (7.52).

Example 8.1: Hard Sphere Scattering
Consider a hard sphere of radius a. Find the contribution to the total cross section from s and p wave scattering as a function of the momentum.
Solution:
The $\ell = 0$ case is simple as the solutions for ingoing and outgoing waves in the region $r > a$ are the Hankel functions which must go to zero at $r = a$.

$$R_{\ell=0}(a) = \frac{1}{2}\left(e^{2i\delta_0}h_0(ka) + h_0^*(ka)\right) = 0$$

Plugging in the expressions for h_0, one obtains,

$$\begin{aligned} 0 &= e^{2i\delta_0}\left(-ie^{ika} + ie^{-ika}\right) \\ &= e^{i\delta_0}\left(-ie^{i(ka+\delta_0)} + ie^{i(ka-\delta_0)}\right) \\ &= 2e^{i\delta_0}\sin(ka+\delta_0). \end{aligned}$$

which gives the $\ell = 0$ phase shifts,

$$\delta_0 = -ka.$$

The contribution to the cross section from the $\ell = 0$ partial waves is then,

$$\sigma_0 = \frac{4\pi}{k^2}\sin^2(ka).$$

As $k \to 0$, the cross section approaches $4\pi a^2$, four times the expected geometric cross section. Calculating the contribution for the $\ell = 1$ partial waves is a bit more difficult. In this case the incoming and outgoing waves are

$$R_{\ell=1}(a) = \frac{1}{2}\left(e^{2i\delta_1}h_1(ka) + h_1^*(ka)\right) = 0.$$

Using Eq.s(8.13) and (8.14),

$$\begin{aligned} h_1(x) &= -\frac{e^{ix}}{x} - i\frac{e^{ix}}{x^2} \\ &= \left[-\frac{\cos(x)}{x} + \frac{\sin(x)}{x^2}\right] + i\left[-\frac{\sin(x)}{x} - \frac{\cos(x)}{x^2}\right], \end{aligned} \tag{8.24}$$

which yields the following expression for δ_1 when requiring that $R_{\ell=1}(a) = 0$,

$$\begin{aligned} e^{2i\delta_1} h_1(ka) + h_1^*(ka) &= 0, \\ e^{2i\delta_1} &= -\frac{h_1^*(ka)}{h_1(ka)} \\ &= \frac{i\left[\cos(ka)/ka - \sin(ka)/(ka)^2\right] + \left[\sin(ka)/ka + \cos(ka)/(ka)^2\right]}{-i\left[\cos(ka)/ka - \sin(ka)/(ka)^2\right] + \left[\sin(ka)/ka + \cos(ka)/(ka)^2\right]} . \end{aligned}$$

This can be solved for δ_1,

$$\tan\delta_1 = \frac{\cos(ka) - \sin(ka)/(ka)}{\sin(ka) + \cos(ka)/(ka)} .$$

The contribution to the total cross section is

$$\sigma_1 = \frac{12\pi}{k^2}\sin^2\delta_1.$$

At low momentum, this becomes

$$\begin{aligned} \delta_1 &\approx \frac{\left[-(ka)^2/2 + (ka)^2/6\right]}{1/ka} \\ &= -\frac{(ka)^3}{3}, \\ \sigma_1 &\approx \frac{4\pi}{3}k^4a^6. \end{aligned} \tag{8.25}$$

In general, for potential scattering, phase shifts tend to behave as $k^{2\ell+1}$ as $k \to 0$, which means that cross sections behave as $k^{4\ell}$. Thus, whereas the cross section for s-wave scattering approaches a constant as $k \to 0$, the p-wave contribution rises very slowly.

8.2 The Optical Theorem

Looking at the expansion of the wave function for large r,

$$\psi_{\vec{k}}(\vec{r})|_{R\to\infty} = e^{i\vec{k}\cdot\vec{r}} + \frac{e^{ikr}}{r}f(\Omega), \tag{8.26}$$

one might ask how the scattered flux, $v|f(\Omega)|^2$, is balanced by a loss of flux in the plane wave. When one squares $\psi_{\vec{k}}$, three terms arise. The first is the squared plane wave, which is unity, regardless of whether scattering occurred. The second part is the flux of the scattered wave, which behaves as $|f(\Omega)|^2$. The third possibility term is the cross term between the plane wave and the scattered wave, which will be linear in $f(\Omega)$. It is this term which must somehow represent the loss of forward going flux.

The scattering amplitude, $f(\Omega)$, defined in Eq. (8.19), is complex. Further ahead, it will be seen

that only the imaginary part of $f(\Omega)$ contributes to the flux. Taking the imaginary part,

$$\begin{aligned}\Im\, f(\Omega) &\equiv \Im \sum_\ell (2\ell+1) e^{i\delta_\ell} \sin\delta_\ell \frac{1}{k} P_\ell(\cos\theta) \\ &= \sum_\ell (2\ell+1)\sin^2\delta_\ell \frac{1}{k} P_\ell(\cos\theta).\end{aligned} \tag{8.27}$$

In the forward direction, $\theta = 0$, $P_\ell(\theta = 0) = 1$ for all ℓ, and

$$\Im\, f(\Omega) = \frac{1}{k}\sum_\ell (2\ell+1)\sin^2\delta, \tag{8.28}$$

and after comparing to expression for the total cross section in Eq. (8.23) becomes

$$\sigma = \frac{4\pi}{k}\Im f(\Omega = 0). \tag{8.29}$$

This is known as the optical theorem.

The forward scattering amplitude is thus related to the dissipation of the plane wave. This was not surprising given the considerations of Sec. 7.9 which showed how the propagator in the forward direction was related to the cross section. To see that this accounts for the missing flux, one can write a term for the flux,

$$\vec{F} = \frac{1}{W}\frac{-i\hbar}{2m}\left\{\psi^*_{\vec{k}}(\vec{r})\nabla\psi_{\vec{k}}(\vec{r}) - (\nabla\psi^*_{\vec{k}}(\vec{r}))\psi_{\vec{k}}(\vec{r})\right\}. \tag{8.30}$$

where W is some arbitrarily large volume. The net rate at which particles enter/leave a spherical surface of radius r enclosing the scattering point is

$$\begin{aligned}\frac{dN}{dt} &= r^2\int d\vec{\Omega}\cdot\vec{F}(\vec{r}) \\ &= \frac{\hbar}{2m}r^2\int d\vec{\Omega}\cdot\vec{k}\,.\end{aligned} \tag{8.31}$$

Taking the gradient, then keeping only those terms which fall slowest in $1/r$,

$$\Delta\vec{F} = \frac{i}{W}\frac{k}{mr}(\hat{r}+\hat{z})\Im f(\Omega)e^{ikr(1-\cos\theta)} + \frac{k}{W}\hat{r}\frac{|f(\Omega)|^2}{mr^2}. \tag{8.32}$$

Here, the flux of the original plane wave, $\hbar\vec{k}/mW$, is subtracted. As $kr\to\infty$,

$$e^{ikr(1-\cos\theta)}\big|_{kr\to\infty} = i\frac{\delta(1-\cos\theta)}{kr}. \tag{8.33}$$

To see that the delta function ensues, one can first see that as long as $\theta \neq 0$ that for large r that the function oscillates infinitely rapidly as a function of k. It can thus be considered to be zero for any wave packet constructed over some small, but non-zero, range of k. But, this oscillation

disappears for $\theta = 0$ and because integrating over $\cos\theta$ gives a constant it must be proportional to a delta function. This means that

$$\Delta\vec{F} = \frac{1}{mWr^2}\left\{-k(\hat{r}+\hat{z})\delta(1-\cos\theta)\Im f(\Omega) + |f(\Omega)|^2\right\}. \tag{8.34}$$

The net rate of particles through the enclosing surface is then

$$\begin{aligned} r^2\int d\vec{\Omega}\cdot\vec{F} &= \frac{1}{mW}\left\{4\pi\Im f(\Omega=0) + k\int d\Omega |f(\Omega)|^2\right\} \\ &= \frac{1}{mW}\left\{4\pi\Im f(\Omega=0) + k\int d\Omega\frac{d\sigma}{d\Omega}\right\} \\ &= \frac{1}{mW}\left\{4\pi\Im f(\Omega=0) + k\sigma\right\}. \end{aligned} \tag{8.35}$$

Thus, if the net flux is to be zero, one must satisfy the optical theorem, Eq. (8.29). This shows that the optical theorem is equivalent to stating that the forward scattered wave must interfere with the original plane wave in such a way that the interference term between the plane wave and the scattered wave results in a contribution to the flux that is equal but opposite to the net flux from the scattered wave.

8.3 Calculating Phase Shifts Numerically

Consider a potential that goes to zero for $r > b$. Rather than considering a solution for the radial wave function $R_\ell(r)$, it is easier to consider a solution for $u_\ell(k,r) = krR_\ell(r)$.

$$\begin{aligned} u_\ell(r) &\equiv krR_\ell(r), \\ \frac{\hbar^2k^2}{2m}u_\ell(r) &= -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2}u_\ell(r) + \frac{\hbar^2\ell(\ell+1)}{2mr^2}u_\ell(r) + V(r)u_\ell(r). \end{aligned} \tag{8.36}$$

Thus, u_ℓ appears to be the solution to a one-dimensional Schrödinger equation with the effective potential,

$$V_{\text{eff}}(r) = V(r) + \frac{\hbar^2\ell(\ell+1)}{2mr^2}, \tag{8.37}$$

and the extra condition that $u_\ell(r\to 0) = 0$.

The solution u_ℓ can be considered as the linear combination of an incoming and phase-shifted outgoing wave.

$$u_\ell(r) = u_\ell^{\text{in}}(r) + e^{2i\delta_\ell}u_\ell^{\text{out}}, \tag{8.38}$$

where the incoming and outgoing waves are known solutions when $r > b$,

$$\begin{aligned} u_\ell^{\text{in}}(r>b) &= krh_\ell^*(r), \\ u_\ell^{\text{out}}(r>b) &= u_\ell^{\text{in},*}(r). \end{aligned} \tag{8.39}$$

To find the phase shift numerically, one can discretize space into steps of Δ, then choose two points r_{n+1} and r_n which are chosen beyond b. One can calculate $u_\ell^{\text{in}}(r_{n+1})$ and $u_\ell^{\text{in}}(r_n)$. One can then numerically solve the discretized Schrödinger equation,

$$-\frac{\hbar^2}{2m\Delta^2}\left(u_\ell^{\text{in}}(r_{n+1}) - 2u_\ell^{\text{in}}(r_n) + u_\ell^{\text{in}}(r_{n-1})\right) = \left(\frac{\hbar^2k^2}{2m} - V_{\text{eff}}(r_n)\right)u_\ell^{\text{in}}(r_n), \tag{8.40}$$

to find $u_\ell^{\rm in}(r_{n-1})$. One can continue iteratively until one finds $u_\ell^{\rm in}(r \to 0)$. The phase shift δ_ℓ is then chosen to make $u_\ell(r \to 0) = 0$. Thus, δ_ℓ is determined by the phase of $u_\ell^{\rm in}$ as $r \to 0$.

$$e^{2i\delta_\ell} = -\left.\frac{u_\ell^{\rm in,*}}{u_\ell^{\rm in}}\right|_{r\to 0} \quad . \tag{8.41}$$

An alternative method is to integrate from $r = 0$, then match logarithmic derivatives at $r = b$ to find the phase shifts. However, this method can be troublesome when it comes to larger ℓ because the wave functions grow as $r^{\ell+1}$ near $r = 0$.

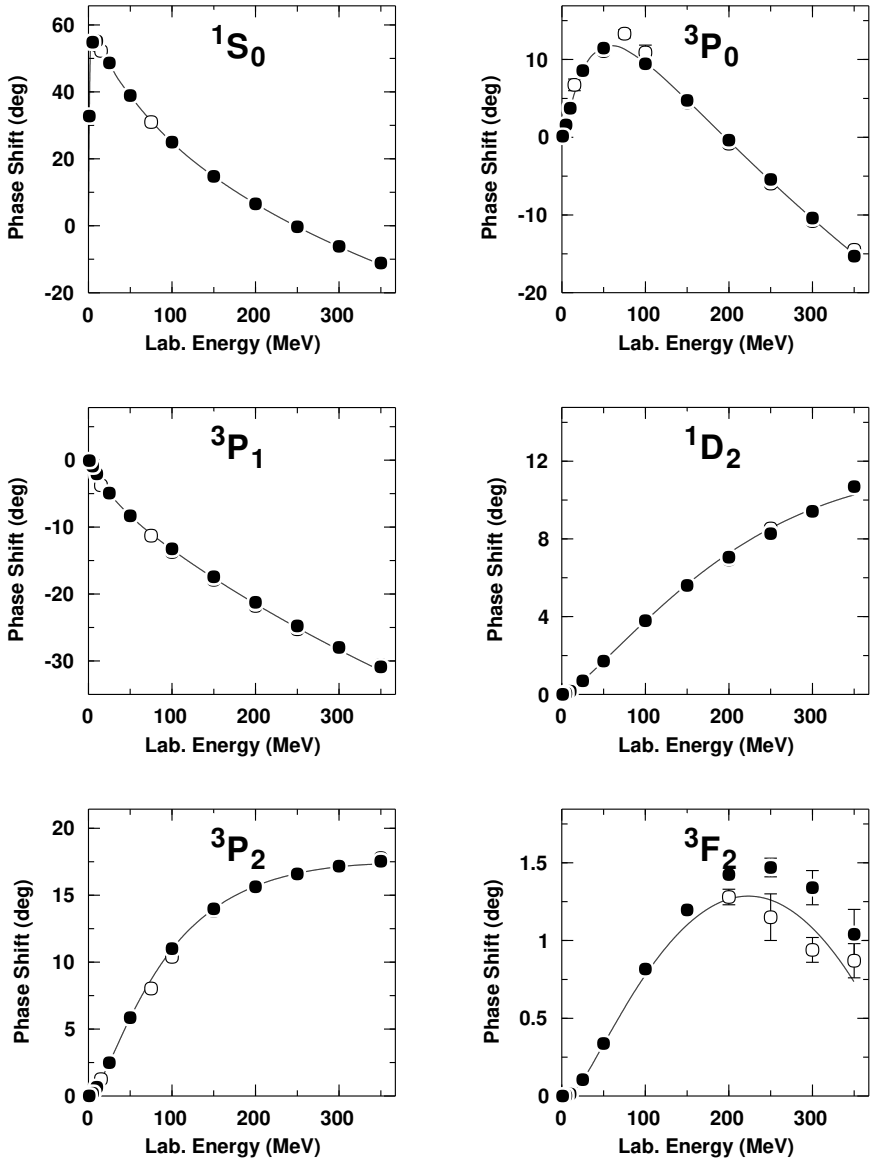
In summary, phase shifts offer a convenient means to express all the information required to describe scattering. In the presence of a spherically symmetric potential, each partial wave becomes a linear combination of the original incoming wave and a phase shifted outgoing wave. Because measurements are made outside the range of the potential the phase shifts, $\delta_\ell(E)$, completely describe all behavior.

One does not usually need to include higher values of ℓ in a phase shift analysis. This is because the partial waves tend to sample distances of order $r > \ell/k$. As $k \to 0$, s-wave scattering dominates the picture. Rarely does one see analyses for $\ell > 3$.

Phase shift analyses are common in low-energy nuclear scattering and in atomic physics. They become rather irrelevant at high energy due to the fact that inelastic channels ($A + B \to C + D + E$) tend to open at high energy, at which point the phase shift vocabulary is insufficient. Even when ℓ is not a good quantum number, e.g. spin-orbit scattering, but where j is a good quantum number, one can still extract phase shifts using polarized beams, although in these cases the vocabulary includes mixing terms.

Finally, as an example we present a plot of proton-proton phase shifts to give an impression of the degree to which these quantities are analyzed. In this figure, the data (circles) are indistin-

guishable from two well-known models (Nijmegan and Bonn).



8.4 The Low Energy Limit

At low energy, one can solve the Schrödinger equation for a given partial wave between $r = 0$ and some point $r = b$ where b is sufficiently large such that the potential is zero. Although the norm of R_ℓ is arbitrary, the logarithmic derivative is completely determined by the potential, the energy, and the point b.

$$\alpha_\ell(k, b) \equiv \left. \frac{dR_\ell(k, r)/dr}{R_\ell(k, r)} \right|_{r=b} \tag{8.42}$$

We will consider $k \approx 0$, one will ignore the energy dependence of α_ℓ by considering a small range of energy.

For $r > b$, $R_\ell(k,r)$ must have the form

$$R_\ell(k,r) \propto \left(h_\ell^*(kr) + e^{2i\delta} h_\ell(kr)\right), \tag{8.43}$$

which means that the logarithmic derivative at $r = b$ becomes

$$\alpha(k,b) = \frac{\partial_r \left(h_\ell^*(kr) + e^{2i\delta_\ell} h_\ell(kr)\right)|_{r=b}}{h_\ell^*(kb) + e^{2i\delta_\ell} h_\ell(kb)}. \tag{8.44}$$

Using the definition of $h_\ell = j_\ell + i n_\ell$ allows one (with some algebra) to write the phase shift in terms of α, j_ℓ and n_ℓ,

$$\cot(\delta_\ell) = \frac{\partial n_\ell/\partial r|_{r=b} - \alpha n_\ell(kb)}{\partial j_\ell(kr)/\partial r|_{r=b} - \alpha j_\ell(kb)}. \tag{8.45}$$

Thus, by finding the logarithmic derivative at $r = b$, one determines the phase shift. Note that if the potential is zero, R_ℓ would be proportional to j_ℓ and the denominator would diverge forcing the phase shift to zero.

We are now in a position to consider the behavior at low k where j_ℓ and n_ℓ have the following behavior,

$$\begin{aligned} j_\ell(kr) &\to \frac{(kr)^\ell}{(2\ell+1)!!} \\ n_\ell(kr) &\to \frac{(\ell-1)!!}{(kr)^{\ell+1}}, \end{aligned} \tag{8.46}$$

where $(2\ell+1)!! \equiv 1 \cdot 3 \cdot 5 \cdots (2\ell+1)$. Inserting these into the expression for the phase shift above.

$$\cot \delta_\ell(k) \approx (kb)^{-(2\ell+1)} (2\ell-1)!!(2\ell+1)!! \frac{\ell + 1 + b\alpha_\ell(k,b)}{\ell - b\alpha_\ell(k,b)}\,. \tag{8.47}$$

For low k, the kinetic term in Schrödinger's equation is negligible compared to the potential and α_ℓ approaches a constant. Thus, the momentum dependence of the phase shifts at low relative momentum is

$$\sin \delta_\ell(k) \propto k^{2\ell+1}. \tag{8.48}$$

One can see that all phase shifts tend to an integral multiple of π at $k = 0$, and that the cross section is dominated by the s-wave contribution at low energy. In fact, the *scattering length*, a, is defined as the derivative of the $\ell = 0$ phase shift at $k = 0$,

$$a \equiv -\frac{\partial}{\partial k} \delta_0(k)|_{k=0}. \tag{8.49}$$

The cross section at very low energy is then,

$$\sigma \approx \frac{4\pi}{k^2} \sin^2(ka) = 4\pi a^2. \tag{8.50}$$

Example 8.2: Scattering off Spherical "Square-Well" Potential
A simple example that is common on exams is that of s-wave scattering off a square well potential. Consider the repulsive potential

$$V(r) = \begin{cases} V_0, & r < b \\ 0, & r > b \end{cases} \tag{8.51}$$

Find the scattering length and the cross section at $k \approx 0$ for a particle of mass m.
We need only consider the s wave in this case. Using the definition $u(k,r) \equiv rR_{\ell=0}(k,r)$, one knows that the Schrödinger equation for $u(k,r)$ looks exactly like a one-dimensional Schrödinger equation. Furthermore, the solution has the following form in the two regions,

$$u_I(k,r) = A\sinh\kappa r, \quad \kappa \equiv \sqrt{\frac{2m(V_0 - E)}{\hbar^2}}, \tag{8.52}$$

$$u_{II}(k,r) = \sin(kr + \delta), \quad k \equiv \sqrt{\frac{2m(E)}{\hbar^2}}$$

Matching logarithmic derivatives at the boundary gives

$$\frac{1}{\kappa}\tanh\kappa b = \frac{1}{k}\tan(kb + \delta). \tag{8.53}$$

Solving for δ for small k,

$$\begin{aligned} \delta &= -kb + \tan^{-1}\left(\frac{k}{\kappa}\tanh\kappa b\right) \\ &\approx k\left(-b + \frac{1}{\kappa}\tanh\kappa b\right) \end{aligned} \tag{8.54}$$

The scattering length and cross section are thus

$$\begin{aligned} a &= b - \frac{1}{\kappa}\tanh\kappa b \\ \sigma(k=0) &= 4\pi a^2. \end{aligned} \tag{8.55}$$

Note that as $V_0 \to \infty$ the phase shift becomes $\delta \to -kb$, which agrees with the result for a hard sphere in Example 8.1. Also, if the potential is small, $\kappa \to 0$, which leads to $a \to 0$ and $\sigma(k=0) \to 0$

8.5 Levinson's Theorem

As shown in the previous section, all phase shifts begin life at multiples of π. If the phase shift at $k = 0$ were anything else an infinite cross section would result at small k because the cross section behaves as $4\pi\sin^2\delta/k^2$. As $k \to \infty$, phase shifts tend to zero because the kinetic energy should overwhelm the potential energy. This asymptotic behavior at large k can be also

understood by realizing that the in the high-energy limit the phase shift is near $V\Delta t/\hbar$, where Δt is the time spent in the potential, which goes to zero as the particle moves very quickly.

Levinson's theorem relates the phase shift at zero energy, which is a multiple of π, to the number of bound states.

$$(2\ell+1)\delta_\ell(k=0) = N_B\pi, \tag{8.56}$$

where N_B is the number of bound states of angular momentum ℓ.

In order to explain the physical motivation of Levinson's theorem we digress to consider the density of states of particles in a large sphere of radius R, which feels a short range potential $V(r)$, with the origin being at the center of the sphere. The wave function at large r is

$$\psi(r\to\infty) \propto \sin(kr+\delta_\ell), \tag{8.57}$$

so that the boundary conditions restrict the possible values of k to

$$kR + \delta_\ell(k) = n\pi. \tag{8.58}$$

Thus, the density of states in momentum is

$$\frac{dn}{dk} = \frac{R}{\pi} + \frac{1}{\pi}\frac{d\delta_\ell}{dk}. \tag{8.59}$$

If one considered $\ell \neq 0$, an additional factor of $(2\ell+1)$ would be added. The change of the density of states due to the non-zero potential is $(2\ell+1)(1/\pi)d\delta_\ell/dk$. The integrated number of extra states inserted between $k=0$ and $k=\infty$ due to the potential is

$$\Delta N_{\text{cont.}} = (2\ell+1)\frac{\delta_\ell(k\to\infty) - \delta_\ell(k=0)}{\pi}. \tag{8.60}$$

However, the net number of states under consideration is not affected by the potential. If $\Delta N_{\text{cont.}}$ states were pushed out of the continuum, then they must have become bound states.

$$N_B + \Delta N_{\text{cont.}} = 0, \tag{8.61}$$

where N_B is the number of bound states. Combining this constraint with the Eq. (8.60) and with the fact that $\delta_\ell(k\to\infty) = 0$ gives Levinson's theorem, Eq. (8.56).

Levinson's theorem is important as it gives one an idea of the general behavior to expect from phase shifts. Attractive potentials tend to have positive phase shifts. If no bound state exists, the phase shifts rise near $k=0$ indicating that the states in the continuum were pulled down toward $k=0$. Then at higher k, the phase shifts fall, indicating that the density of continuum states was depleted at higher k. If a bound state exists, the phase shift would generally start at π and usually fall as a function of k. The falling phase shift denotes a negative correction to the density of states. These are the states from which the bound state was formed.

Phase shifts behave rather peculiarly when the potentials are at the threshold of creating a bound state. A small change in the potential causes the $\delta(k=0)$ to jump from zero to π. In these instances scattering lengths can be anomalously long. Such an example is neutron-neutron scattering. The scattering length is nearly -20 fm, which gives a cross section over 1000 fm^2 even though the range of the potential is only of the order of one fm. Thus, the neutron-neutron cross section at low momentum is ~ 100 times larger than πR^2, where R is the range of the strong interaction.

8.6 Effective Range Theory for $\ell = 0$

The mathematics of low-energy scattering theory is often referred to as effective range theory, and goes back to the work of Schwinger and of Hans Bethe in 1949, https://journals.aps.org/pr/abstract/10.1103/PhysRev.76.38, and was mainly applied to nucleon-nucleon scattering. As $k \to 0$, scattering is dominated by the $\ell = 0$ partial wave. Phase shifts (modulo integer numbers of π) at small momentum rise with powers of $k^{2\ell+1}$, so cross sections, which behave as $(4\pi \sin^2 \delta)/k^2$, behave as $k^{4\ell}$ at low momentum, and for many experiments only the $s-$waves contribute. Effective range theory parameterizes how the $\ell = 0$ cross sections are principally decided by two properties of the potential, the scattering length and the effective range. Given that we are interested in the low-momentum limit, and are thus interested in viewing something akin to a Taylor expansion of the phase shift with respect to the momentum, it is not surprising that two properties of the potential, e.g. the width and depth, would determine the first two terms in the Taylor expansion. However, the effective range expansion, which is an expansion of $\cot \delta$, tends to do remarkably well as the higher-order terms play little role until the momenta exceed the inverse scale of the potential.

The principal relation of effective range theory is derived as follows. First consider the radial solution for the $\ell = 0$ partial wave, $u(k, r) = rR_{\ell=0}(k, r)$. If the potential vanishes for $r > R$, the solution for $r > R$ is of the form,

$$u(k, r > R) = \sin(kr + \delta(k))/Z(k). \tag{8.62}$$

We will pick real solutions (which you can do if the potential is real) and leave off the factor $e^{i\delta}$ from Eq. (8.16). We will also consider a solution to the Schroödinger equation, $w(k, r)$, without the potentil, that matches $u(k, r)$ for $r > R$, i.e., $w(k, r > R) = u(k, r)$. This solution will not satisfy the boundary condition at the origin, i.e.,

$$w(k, r > 0) = \sin(kr + \delta(k))/Z(k). \tag{8.63}$$

Further, we will define the arbitary normalization constant $Z(k) = \sin \delta(k)$. With this definition, $w(k = 0) = 1$. Now, considering solutions u_1 and u_2 at two momenta, k_1 and k_2 respectively, one can consider the Schrödinger equations,

$$\begin{aligned} k_1^2 u_1 &= -\partial_r^2 u_1 + 2m\hbar^2 V(r) u_1, \\ k_2^2 u_2 &= -\partial_r^2 u_2 + 2m\hbar^2 V(r) u_2, \\ k_1^2 w_1 &= -\partial_r^2 w_1, \\ k_2^2 w_2 &= -\partial_r^2 w_2. \end{aligned} \tag{8.64}$$

Then, taking the products, $u_1 u_2$ and $w_1 w_2$,

$$\begin{aligned} (k_1^2 - k_2^2) \int_0^R dr\, (u_1 u_2 - w_1 w_2) \\ = \int_0^R dr\, \left[-(\partial_r^2 u_1) u_2 + (\partial_r^2 u_2) u_1 - (\partial_r^2 w_1) w_2 + (\partial_r^2 w_2) w_1 \right]. \end{aligned} \tag{8.65}$$

The potential $V(r)$ has fallen out of the expression, and be recognizing that $(\partial_r^2 u_1)u_2 - u_1 \partial_r^2 u_2 = \partial_r[(\partial_r u_1)u_2 - u_1 \partial_r u_2]$ is a total derivative, one can perform the integration,

$$(u_1 \partial_r u_2 - u_2 \partial_r u_1 + w_1 \partial_r w_2 - w_2 \partial_r w_1)_{r=0} = (k_1^2 - k_2^2) \int_0^R dr\, (u_1 u_2 - w_1 w_2). \tag{8.66}$$

This expression is exact and applies for any momenta k_1 and k_2. The upper limit of the integration at R vanished because $w_i = u_i$ for $r > R$. Now, because $u_1(r = 0) = u_2(r = 0)$ and $w_i(r = 0) = \sin(\delta_i)$, the expression simplifies further,

$$\begin{aligned}(w_1 \partial_r w_2 - w_2 \partial_r w_1)_{r=0} &= (k_1^2 - k_2^2) \int_0^R dr\ (u_1 u_2 - w_1 w_2), \\ \frac{k_2 \sin(\delta_1)\cos(\delta_2) - k_1 \cos(\delta_1)\sin(\delta_2)}{\sin(\delta_1)\sin(\delta_2)} &= (k_1^2 - k_2^2) \int_0^R dr\ (u_1 u_2 - w_1 w_2). \end{aligned} \tag{8.67}$$

This last expression forms the basis for understanding the low-momentum expansion. For example, taking the limit $k_1 \to 0$, and setting $k_2 = k$,

$$\begin{aligned} k \cot(\delta) &= -\frac{1}{a} + \frac{1}{2}\rho k^2 \\ \rho &\equiv 2 \int_0^R dr\ \left[w(k=0,r) w(k,r) - u(k=0,r) u(k,r) \right], \\ a &\equiv -\left. \frac{d\delta}{dk} \right|_{k=0}. \end{aligned} \tag{8.68}$$

This is the effective range formula. The quantity ρ is known as the effective range. Because $u(r = 0) = 0$ and $w(r = 0) = 1$, it provides a characteristic distance over which the u returns to the asymptotic form, and should be of the length scale of the potential. The scattering length a was defined earlier. For small momenta, the cross section is

$$\sigma = \frac{4\pi}{k^2} \sin^2 \delta \approx 4\pi a^2. \tag{8.69}$$

The scattering length need not be of the range of the potential. In fact, it can be arbitrarily large in the limit that the potential is adjusted to very nearly providing a bound state. For higher momenta terms of higher power of k contribute to the effective range expansion. The scale at which they become important is given by the effective range ρ, i.e., Eq. (8.68) above should be accurate for $k\rho < 1$.

8.7 Coulomb Waves and Gamow Factors

The Coulomb potential is not short-range, and even the incoming wave is altered by the potential at large distance. Hence, the incoming and outgoing spherical waves are not Hankel functions but are instead Coulomb waves. At large r the spherical Coulomb waves behave as

$$R(r) \sim \frac{1}{r} e^{\pm i(kr - \gamma \log r)}, \tag{8.70}$$

where

$$\gamma \equiv \frac{\mu Z_1 Z_2 e^2}{\hbar^2 k} = -\frac{1}{a_0 k}. \tag{8.71}$$

When the potential is attractive, e.g. the Hydrogen atom, γ is negative. The factor γ is known as the Sommerfeld parameter.

The entire solutions for Coulomb waves are known as confluent hypergeometric functions which can be found in books of special functions such as https://en.wikipedia.org/wiki/Abramowitz_and_Stegun. They are also often simply called Coulomb wave functions.

When an extra potential is added to the problem, one can still classify the behavior in terms of phase shifts, but the problem becomes one of phase-shifted Coulomb waves rather than phase-shifted Hankel functions. Of course, all the relations for total cross sections are modified because the Coulomb force produces a scattering all it's own.

The expression for a plane wave plus it's scattered portion is written in terms of hypergeometric functions as

$$\psi_k(r) = \Gamma(1+i\gamma)e^{-\pi\gamma/2}e^{ikz}{}_1F_1(-i\gamma;1;ik(r-z)). \tag{8.72}$$

For small r ($r << a_0$) the hypergeometric function ${}_1F_1$ goes to unity and the solution is approximately,

$$\begin{aligned}\psi_k(r << a_0) &= \Gamma(1+i\gamma)e^{-\pi\gamma/2}e^{ikz} \\ |\psi_k(r << a_0)|^2 &= |\Gamma(1+i\gamma)|^2 e^{-\pi\gamma} \\ &= \frac{2\pi\gamma}{e^{2\pi\gamma}-1}.\end{aligned} \tag{8.73}$$

This last factor is referred to as the Gamow factor or penetrability.

Note that the Gamow factor is independent of r. For repulsive potentials, it can be thought of as the penetration probability of reaching the origin, relative to what it would have been if there were no Coulomb. Of course, in classical physics a charged particle never reaches the origin when there is a repulsive Coulomb barrier. If the potential is repulsive, gamma is positive. As $k \to 0$, $\gamma \to \infty$ and the Gamow factor goes to zero, meaning that low energy waves have great difficulty penetrating the barrier. As k is increased and approaches $(2\pi a_0)^{-1}$ the penetrability rises. For attractive potentials, the Gamow factor is greater than unity. In both the attractive and repulsive cases, the Gamow factor approaches unity as $k \to \infty$.

The Gamow penetrability plays an important role in astrophysical rates. Fusion reactions would happen much more quickly if it were not for the Coulomb barrier. Rather than listing cross sections, reactions are often described by what is known as the astrophysical S factor,

$$S(E) = Ee^{2\pi\gamma}\sigma(E). \tag{8.74}$$

For large barriers this cancels the most divergent part of the Gamow factor, and the resulting quantity, $S(E)$ is well behaved as $E \to 0$.

Example 8.3: When Does a Potential Become Long Range
If a potential falls as α/r^m, for what values of m can one define the scattering with phase shifted partial waves?
Solution:
Many forms for the potential never really go to zero at large r, e.g. the Yukawa form $e^{-\alpha r}/r$, but they do approach zero well enough. For the power-law form considered here, one can realize that for large r the potential falls sufficiently slowly to consider the estimate the contribution to the phase shift from the potential outside some large distance R with the WKB

approximation,

$$\Delta\delta(p) = \int_R \infty dr\ \Delta p(r)/\hbar.$$

Here, Δ denotes that this is the contribution to the phase shift solely from the potential outside R. The momentum $p(r)$ is a function of the potential, $p(r) = \sqrt{2m(E - V(r))}$, and expanding for small V,

$$\Delta\delta(p) \int_R \infty dr\ \sqrt{2mE}\frac{V(r)}{2E}.$$

If the potential falls of more quickly than $1/r$, the integral vanishes for sufficiently large R. However, for the Coulomb potential, the integral is undefined for any R. Thus, the answer is that phase shifted Hankel waves can be considered as long as $m > 1$. One can also see that for any exponentially falling potential, such as the Yukawa form, one can also apply the usual scattering theory.

8.8 Classical Interpretation of the Scattering Wave Function

The scattered state wave function, $\psi(\vec{k}, \vec{r})$, normalized so that the incoming wave is $e^{i\vec{k}\cdot\vec{r}}$, can be thought of as a measure of the relative probability density (after being squared) for a particle of incoming momentum $\vec{k}$ to be found at $\vec{r}$. The solution includes outgoing momenta distributed over all directions. If one takes the complex conjugate of the scattered wave function, one finds a solution where outgoing parts of the solution match a plane wave of momentum $\vec{k}$. The incoming components then include the complex conjugate of the outgoing wave in scattering theory and describe the incoming flux from all directions, that are joined together precisely to form an outgoing wave that matches a plane wave.

One insightful relation is that of the wave function normalization compared to the change in the density of states in terms of $(1/\pi)d\delta/dk$. This can be derived in a similar approach to find the effective range formula above. For a given partial wave,

$$\int dr \left(|u_\ell(r)|^2 - |u_\ell^{(0)}|^2\right) = \frac{1}{2}\frac{d\delta_\ell}{dk}, \tag{8.75}$$

where $u_\ell(k, r) = rR_\ell(k, r)$ and $u_\ell^{(0)}(k, r) = krj_\ell(kr)$ are the solutions for partial waves with and without potentials. This provides insight into the meaning of the squared scattered wave function. It describes the strength with which one can emit a particle into a given outgoing momentum state. One can equivalently think of the fact that a particle with asymptotic momentum $\vec{k}$ is more or less likely to be in the region of the potential due to the presence of the potential, or one can think of being more likely to be in the region of the potential because the density of states changed.

The classical analog to the squared scattered wave function is in terms of the ratio of the phase space at the point $\vec{r}$ compared to the the size of the same phase space element in its asymptotic momentum state. If one emits a particle at position $\vec{r}_0$ into some differential momentum volume

d^3p_0 centered around $\vec{p}_0$, the potential will alter the momentum into some differential volume d^3p_f centered about $\vec{p}_f$. In the limit that the momenta are larger than characteristic inverse sizes of the potential, $p >> \hbar/R$, the wave functions become

$$|\phi(\vec{p},\vec{r})|^2 \rightarrow \frac{d^3p_0}{d^3p_f}. \tag{8.76}$$

This latter ratio is a function of $\vec{r}_0$ and $\vec{p}_f$, and can be found by solving for the classical trajectories, $\vec{p}_0(\vec{p}_f,\vec{r}_0)$. For the case of the Coulomb interaction, this ratio can be found analytically, https://journals.aps.org/prc/pdf/10.1103/PhysRevC.45.387, using energy conservation, angular momentum conservation, and conservation of the Lenz vector.

8.9 Exercises

1. Show that if the function $u_\ell(kr)$ is defined in terms of $R_\ell(r)$

$$u_\ell(kr) \equiv rR_\ell(r),$$

where R_ℓ is a solution to the radial Schrödinger equation

$$\left\{-\frac{\hbar^2}{2m}\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{\hbar^2}{2m}\frac{\ell(\ell+1)}{r^2} + V(r)\right\}R_\ell(r) = \frac{\hbar^2k^2}{2m}R_\ell(r),$$

that u_ℓ satisfies the differential equation,

$$\left(\frac{d^2}{dx^2}+1\right)u_\ell(x) = \frac{\ell(\ell+1)}{x^2}u_\ell(x) + \beta(x)u_\ell(x),$$

where β is proportional to the potential,

$$\beta(x) = \frac{2m}{\hbar^2k^2}V(x/k).$$

2. Recurrence relations for Bessel functions provide you the ability to find forms for solutions at higher ℓ given you know the form for $\ell = 0$ and $\ell = 2$

   (a) Show that in the case of zero potential that the solutions u_ℓ satisfy the recurrence relation.

$$u_{\ell+1}(x) = \frac{(\ell+1)}{x}u_\ell(x) - \frac{d}{dx}u_\ell(x).$$

   Use the expressions from the previous problem,

$$\left(\frac{d^2}{dx^2}+1\right)u_\ell(x) = \frac{\ell(\ell+1)}{x^2}u_\ell(x) + \beta(x)u_\ell(x). \quad (1)$$

   (b) Show that this recurrence relation can be equivalently expressed as

$$f_{\ell+1}(x) = \frac{\ell}{x}f_\ell(x) - \frac{d}{dx}f_\ell(x),$$

   where f_ℓ is a solution to the radial Schrödinger equation, $f_\ell(kr) \equiv u_\ell(kr)/(kr)$, which means that f_ℓ might be any linear combination of j_ℓ and n_ℓ.

(c) One can also show that a second recurrence relation is satisfied,

$$f_{\ell-1}(x) = \frac{(\ell+1)}{x} f_\ell(x) + \frac{d}{dx} f_\ell(x).$$

Given this recurrence relation, plus the one from the previous problem, show that

$$f_{\ell-1}(x) + f_{\ell+1}(x) = \frac{(2\ell+1)}{x} f_\ell(x)$$

(d) Using expressions for j_0, j_1, n_0 and n_1, use recurrence relations to find expressions for j_2 and n_2.

(e) Using the recurrence relations, show that $j_\ell(z)$ and $n_\ell(z)$ behave as z^ℓ and $z^{-(\ell+1)}$ respectively for $z \to 0$. Begin with the facts that $j_0(z)$ and $n_0(z)$ behave as z^0 and z^{-1} respectively, and that they are even and odd functions in z.

3. Consider a particle of mass m that interacts with a spherically symmetric attractive potential.

$$V(r) = \begin{cases} -V_0, & r < b \\ 0, & r > b \end{cases}$$

(a) What is the minimum depth V_{min} that allows a bound state?

(b) Find an expression for the phase shift in terms of a particle whose momentum is p.

(c) Assuming the depth is $V_0 = 0.99 \cdot V_{\text{min}}$, plot the s-wave phase shift for momenta in the range $0 < p < 5\hbar/b$. Use units of $\hbar/b$ for the momenta.

(d) Repeat the above problem for $V_0 = 1.01 \cdot V_{\text{min}}$.

(e) What are the scattering lengths for the two potentials?

4. Consider a proton scattering off of a an attractive one-dimensional potential,

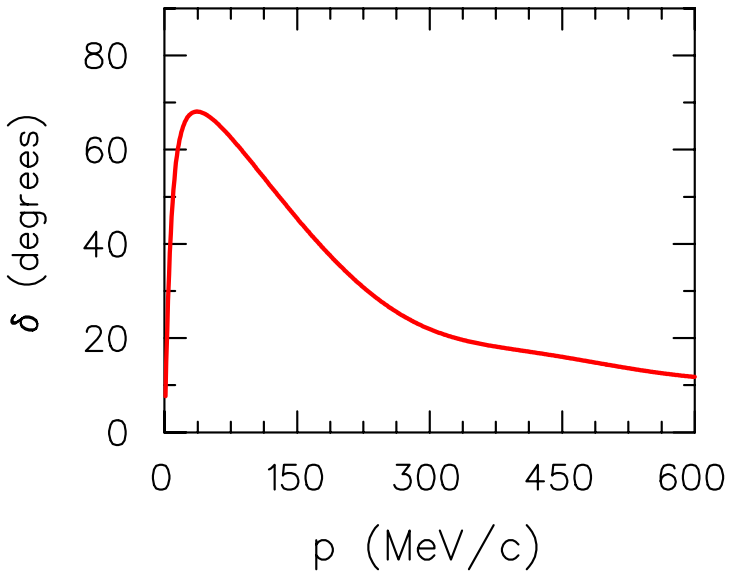
$$V(x) = \begin{cases} \infty, & x < 0 \\ -V_0\left(1 - \frac{r^2}{R^2}\right), & 0 < x < R \\ 0, & r > R \end{cases}$$

For this example, we will consider $R = 2.5$ fm, and V_0=16 MeV. If you wish, to make the units more natural, you may consider $\hbar c$=197.327 MeV·fm, and $m_p = 938.27$ MeV/c^2. Consider a particle incident on the well with energy E that enters and leaves the well with energy E. Far away, the solutions are of the form,

$$\psi(x) = e^{-ipx/\hbar} - e^{2i\delta + ipx/\hbar}, \quad x >> R$$

(a) Programming in either PYTHON or C++, construct a program that runs and returns a listing of δ vs. p for $0 < p < 600$ MeV/c, in steps of 2.0 MeV/c.

A graph of the results:



(b) **EXTRA CREDIT Make a graph like the one above, except for the region between p=0 and p=1.0 MeV, and consider two strengths of the potential, $V_0 = 17.0$ MeV and $V_0 = 17.025$ MeV. Be sure to calculate values for very small values of p, in steps of .001 MeV. For this problem, turn in a paper copy of the graph.**

5. Consider a potential which gives non-zero phase shifts for $0 \leq \ell \leq \ell_{\max}$, where $\ell_{\max}$ is a large number. Assume these phase shifts can be considered as random numbers, evenly distributed between zero and 2π. Using the expression for the cross section,

$$\sigma = \frac{4\pi\hbar^2}{p^2} \sum_{\ell} (2\ell + 1) \sin^2 \delta_\ell,$$

(a) Find the overall cross section by averaging over the expectation of the random phases. Give your answer in terms of $\ell_{\max}$ and the incoming momentum p.

(b) Consider a problem classically where one scatters off a strong central potential whose maximum range is $R_{\max}$. From classical arguments, what is the maximum angular momentum of a particle that scatters? Give your answer in terms of $R_{\max}$ and the incoming momentum p. What is the total cross section in terms of $R_{\max}$ in the limit that $\ell_{\max}$ is large.

6. A particle of mass m experiences an attractive spherically symmetric potential,

$$V(r) = -\beta\delta(r - a),$$

where $\beta > 0$.

(a) In terms of a, and the electron mass m, what is the minimum value of β that results in a bound state?

(b) What is the scattering length and the cross section in the limit that the incident beam energy is zero.

(c) If a scattered wave in a large volume behaves as

$$\psi(\vec{k},\vec{r},t) \sim e^{i\vec{k}\cdot\vec{r}-i\omega t},\ t\to\infty$$

in the outgoing limit (large time after interacting with potential), what is the relative probability,

$$\alpha(k) = \frac{\rho(\vec{r}=0)}{\rho_0(\vec{r}=0)},$$

that it will appear at the origin while interacting with the potential? Here ρ_0 is the probability density (per unit volume) in the absence of the potential, and ρ is the probability density with the potential in place. FYI: The ratio α would be the same if the boundary conditions specified an incoming plane wave, instead of matching to an outgoing plane wave.

(d) Assume β is sufficiently large to bind a particle, and that the ground state energy is $-B$. For the ground state what is the probability density of finding the particle at $\vec{r}=0$? Refer to this as $\rho_b(\vec{r}=0)$? Given answer in terms of a and the binding energy B (or equivalently the decay wave number, $q \equiv \sqrt{2mB/\hbar^2}$). HINT: You don't need to solve for the binding energy!

7. Near a resonance of energy ϵ_R, a phase shift behaves as:

$$\tan\delta_\ell = \frac{\Gamma/2}{\epsilon_R - E},$$

where E is the c.m. kinetic energy. For the following problems, assume that $\Gamma << \epsilon_R$, so that the $4\pi/k^2$ prefactor in the expression for the cross section can be considered as a constant.

(a) Write down the cross section $\sigma_\ell(E)$.

(b) What is the maximum cross section for a narrow cross section (as E is varied) for scattering through that partial wave? (How does it depend on ϵ_R, Γ, the reduced mass μ, and ℓ)?

(c) What is the energy integrated cross section ($\int \sigma_\ell(E)dE$)?

8. The temperature at the center of the sun is 15 million degrees Kelvin. Consider two protons with a relative kinetic energy characteristic of the temperature,

$$\frac{\hbar^2k^2}{2\mu} = \frac{3}{2}kT.$$

(a) What is the Gamow penetrability factor? Give a numeric value.

(b) If the two particles were a proton and a ^{12}C nucleus, what would the penetrability factor become?

9. Consider a particle of mass m undergoing a repulsive spherically symmetric Coulomb potential, $V = Ze^2/r$. The classical analogue of the squared wave function is

$$|\phi(\vec{p}_f,\vec{r})|^2 \to \frac{d^3p_i}{d^3p_f}.$$

Here, $\vec{p}_i$ is the momentum when the particle is at position $\vec{r}$, and $\vec{p}_f$ is the asymptotic momentum at large times.

(a) If one averages over all directions of the final momentum, what is $\langle|\phi(\vec{p}_f, \vec{r})|^2\rangle$? Give a sketch of the classical approximationn to $\langle|\phi(\vec{p}_f, \vec{r})|\rangle$ as a function of r for fixed p.

(b) Repeats (a) but working in two dimensions, i.e. find d^2p_i/d^2p_f.

(c) Repeats (a) and (b) but working in one dimension, i.e. find dp_i/dp_f.

9 Second Quantization and Radiation

9.1 Quantum Fields and Second Quantization

Thus far, our study of wave functions, $\psi(\vec{r})$ in coordinate space or ψ_k in momentums space, has considered the quantum mechanics of a single particle. The particle exists in some kind of state or at some position, and exists in perpetuity – at any time there is a single particle. In reality, particles are created and destroyed and the net particle number is in constant flux. The most obvious example is the emission or absorption of a photon. But even charged particles are constantly being created in particle-antiparticle pairs. For example, the eigentstate referred to a as photon includes contributions from electron-antielectron pairs. Even the vacuum includes the creation of such pairs, though the pairs exist only for a short time determined by the $\boldsymbol{\Delta E \Delta t \gtrsim \hbar/2}$ uncertainty principle. In addition to electromagnetic decays, most other decays in particle physics involve the creation of new particles such as a Higgs boson decaying to a bottom and anti-bottom quark. In relativistic quantum field theory, all interactions, e.g. the Coulomb force, are enacted through the creation and absorption of particles. The potential terms in Hamiltonians involve operators that create or destroy particles, with such operators sometimes being referred to a second quantization. For this class, we will focus our study of second quantization on decays.

Creation and destruction operators were introduced to describe the creation of energy quanta in a harmonic oscillator. The operators obeyed the relations,

$$[a_i, a_j^\dagger] = \delta_{ij}, \;\; [a_i, a_j] = 0, \;\; [a_i^\dagger, a_j^\dagger] = 0, \tag{9.1}$$

where i specifies which oscillator is being affected. If one had N harmonic oscillators, one would have N independent creation operators and N independent destruction operators. Note that the operators corresponding to different oscillators commute with one another as they are unrelated.

The essential feature of creation and destruction operators is that they increase/decrease the number of quanta, where the number of quanta of the oscillator i is found by using the operator $N_i = a_i^\dagger a_i$. However, instead of counting only energy quanta, as in the case of the harmonic oscillator, the number operator could also refer to a number of particles in some single-particle level i. Let us then consider a creation operator, $a_k^\dagger$, for each mode, k, where a mode corresponds to a single-particle eigenstate. The one-particle state, $|k\rangle$, of momentum k, is created by $a_k^\dagger$ operating on the vacuum.

$$|k\rangle = a_k^\dagger |0\rangle. \tag{9.2}$$

The state is normalized to unity, just as one would expect for creation operators.

$$\langle k'|k\rangle = \delta_{k'k}. \tag{9.3}$$

By operating twice with $a_k^\dagger$, one creates a state with two particles of momentum k. Because the mode k might have any number of such particles the complete set of states for a single mode would include $|0\rangle, a_k^\dagger|0\rangle, \cdots, [(a_k^\dagger)^n/\sqrt{n!}]|0\rangle, \cdots$.

Example 9.1: Coupled Harmonic Oscillators
Here, we consider two coupled harmonic oscillators. The oscillators are described by creation operators $a_1^\dagger$ and $a_2^\dagger$. Let the Hamiltonian be

$$\begin{aligned} H &= H_0 + V, \\ H_0 &= \epsilon_1 a_1^\dagger a_1 + \epsilon_2 a_2^\dagger a_2, \\ V &= \beta \left(a_2^\dagger a_1 + a_1^\dagger a_2 \right). \end{aligned} \tag{9.4}$$

Consider the operators $b_1^\dagger$ and $b_2^\dagger$ defined by

$$\begin{aligned} b_1^\dagger &\equiv a_1^\dagger \cos\theta + a_2^\dagger \sin\theta, \\ b_2^\dagger &\equiv a_2^\dagger \cos\theta - a_1^\dagger \sin\theta. \end{aligned} \tag{9.5}$$

a) Show that b_1, b_2, $b_1^\dagger$ and $b_2^\dagger$ obey the commutation rules for destruction operators.

b) Find E_1, E_2 and θ so that

$$H = E_1 b_1^\dagger b_1 + E_2 b_2^\dagger b_2.$$

Solution:
(a) To test that b_i and $b_i^\dagger$ act as creation operators,

$$\begin{aligned} [b_1, b_1^\dagger] &= [(a_1 \cos\theta + a_2 \sin\theta), (a_1^\dagger \cos\theta + a_2^\dagger \sin\theta)] \\ &= \cos^2\theta + \sin^2\theta = 1, \\ [b_2, b_2^\dagger] &= [(a_2 \cos\theta - a_1 \sin\theta), (a_2^\dagger \cos\theta - a_1^\dagger \sin\theta)] \\ &= \cos^2\theta + \sin^2\theta = 1, \\ [b_1, b_2^\dagger] &= [(a_1 \cos\theta + a_2 \sin\theta), (a_2^\dagger \cos\theta - a_1^\dagger \sin\theta)] \\ &= -\sin\theta\cos\theta + \sin\theta\cos\theta = 0, \\ [b_1, b_1] &= [b_2, b_2] = [b_1, b_2] = 0. \end{aligned}$$

By taking the complex conjugates, one can see that $[b_1^\dagger, b_1^\dagger] = [b_2^\dagger, b_2^\dagger] = [b_1^\dagger, b_2^\dagger] = 0$.
(b) First, expand the Hamiltonian in the desired form in terms of a_i and $a_i^\dagger$ using Eq.s (9.5),

$$\begin{aligned} H &= E_1 b_1^\dagger b_1 + E_2 b_2^\dagger b_2 \\ &= \frac{E_1 + E_2}{2} \left(a_1^\dagger a_1 + a_2^\dagger a_2 \right) + \frac{E_1 - E_2}{2} \left(a_1^\dagger a_1 - a_2^\dagger a_2 \right) \cos 2\theta \\ &+ \frac{E_1 - E_2}{2} \left(a_1^\dagger a_2 + a_2^\dagger a_1 \right) \sin 2\theta, \end{aligned}$$

then compare it to the original Hamiltonian in Eq. (9.4),

$$H = \frac{\epsilon_1 + \epsilon_2}{2} \left(a_1^\dagger a_1 + a_2^\dagger a_2 \right) + \frac{\epsilon_1 - \epsilon_2}{2} \left(a_1^\dagger a_1 - a_2^\dagger a_2 \right) + \beta \left(a_1^\dagger a_2 + a_2^\dagger a_1 \right).$$

To determine E_1, E_2 and β set the two Hamiltonians to be the same,

$$\begin{aligned}\epsilon_1 + \epsilon_2 &= E_1 + E_2,\\ \epsilon_1 - \epsilon_2 &= (E_1 - E_2)\cos 2\theta,\\ \beta &= \frac{1}{2}(E_1 - E_2)\sin 2\theta.\end{aligned}$$

The Hamiltonians are equivalent when

$$\begin{aligned}\tan 2\theta &= \frac{2\beta}{\epsilon_1 - \epsilon_2},\\ E_1 + E_2 &= (\epsilon_1 + \epsilon_2),\\ \left(\frac{E_1 - E_2}{2}\right)^2 &= \beta^2 + \left(\frac{\epsilon_1 - \epsilon_2}{2}\right)^2.\end{aligned}$$

One can note the algebraic equivalence of this problem to the two-component problem with a Hamiltonian

$$H = \frac{\epsilon_1 + \epsilon_2}{2}\mathbb{I} + \frac{\epsilon_1 - \epsilon_2}{2}\sigma_z + \beta\sigma_x\,,$$

where one is asked to find the eigenvalues E_1 and E_2.

9.2 Field Operators: Creation and Destruction Operators in Coordinate Space

The operator $a^\dagger_{\vec{k}}$ will create a particle with momentum $\vec{k}$, where the momentum states are assumed to be discrete. The operator $a_{\vec{k}}$ will remove a particle from such a state. The particle might be even be massive or charged. If the state $|0\rangle$ is the vacuum, $a^\dagger_{\vec{k}}|0\rangle$ is a state with one particle in momentum state $\vec{k}$. Field operators are the coordinate-space analogs to $a_{\vec{k}}$ and $a^\dagger_{\vec{k}}$. They are defined as

$$\begin{aligned}\Psi(\vec{x}) &= \sum_{\vec{k}} a_{\vec{k}}\frac{e^{i\vec{k}\cdot\vec{x}}}{\sqrt{V}},\\ \Psi^\dagger(\vec{x}) &= \sum_{\vec{k}} a^\dagger_{\vec{k}}\frac{e^{-i\vec{k}\cdot\vec{x}}}{\sqrt{V}}.\end{aligned} \tag{9.6}$$

They obey the commutation rules,

$$\begin{aligned}[\Psi(\vec{x}), \Psi^\dagger(\vec{y})] &= \frac{1}{V}\sum_{\vec{k},\vec{k}'} e^{i(\vec{k}\cdot\vec{x}-\vec{k}'\cdot\vec{y})}[a_{\vec{k}}, a^\dagger_{\vec{k}'}] \\ &= \frac{1}{V}\sum_{\vec{k}} e^{i\vec{k}\cdot(\vec{x}-\vec{y})} \\ &= \frac{1}{(2\pi)^3}\int d^3k\, e^{i\vec{k}\cdot(\vec{x}-\vec{y})} \\ &= \delta^3(\vec{x}-\vec{y}).\end{aligned} \tag{9.7}$$

Further, they return zero when acting on the vacuum,

$$\Psi(\vec{x})|0\rangle = 0, \quad \langle 0|\Psi^\dagger(\vec{x}) = 0. \tag{9.8}$$

These operators create the state $|\vec{x}\rangle$,

$$|\vec{x}\rangle = \Psi^\dagger(\vec{x})|0\rangle, \;\; \langle\vec{x}| = \langle 0|\Psi(\vec{x}). \tag{9.9}$$

Here, $|\vec{x}\rangle$ is a state with one particle at the position $\vec{x}$, and is normalized as

$$\langle\vec{x}|\vec{y}\rangle = \langle 0|\Psi(x)\Psi^\dagger(\vec{y})|0\rangle = \delta^3(\vec{x}-\vec{y}). \tag{9.10}$$

One can easily check that

$$\langle\vec{x}|\vec{k}\rangle = \langle 0|\Psi(\vec{x})a^\dagger_{\vec{k}}|0\rangle = \frac{e^{i\vec{k}\cdot\vec{x}}}{\sqrt{V}}. \tag{9.11}$$

One should keep in mind that $\Psi(\vec{x})$ is an operator, not a wave function. To remind the reader of the difference, a wave function, $\langle\vec{x}|\psi\rangle = \psi(\vec{x})$, is an expectation value and is a complex number. It includes a bra and ket. For the case of a finite number of states, a bra or ket is a vector and an expectation has the form $\langle\chi|A|\phi\rangle = \chi_i^* A_{ij}\phi_j$ and there are no remaining indices. In contrast an operator is a matrix, e.g., A_{ij}. Each ij element of the operator projects out the j^{th} element of the vector described by the ket, then transforms it to $|i\rangle$. The creation operator $\Psi^\dagger(\vec{x})$ operating on some state ϕ replaces ϕ with the same state, but with a particle added at $\vec{x}$. Unlike the cases from the previous chapters the position labels $\vec{x}$ or the momentum labels $\vec{p}$ are insufficient to describe a state. A complete basis must include configurations with different numbers of particles. Thus, one needs operators that describe the change from states with different particle number. Creation and destruction operators provide a means to describe elements involving n and $n \pm 1$ particles. A product of creation or destruction operators is required to describe a transition between states which differ by more than 2 in particle number. The fact that much of physics, including the standard model of particle physics, is modeled by a Hamiltonian with powers of creation operators no higher than the quartic level is indeed a profound property of nature.

If $|\phi\rangle$ refers to a one-particle state, the state is related to the wave function $\phi(x)$ by

$$\begin{aligned} |\phi\rangle &= \int d^3x\, \phi(x)\Psi^\dagger(\vec{x})|0\rangle, \\ \Psi(\vec{x})|\phi\rangle &= \int d^3x'\, \phi(\vec{x}')\Psi(\vec{x})\Psi^\dagger(\vec{x}')|0\rangle \\ &= \phi(\vec{x})|0\rangle, \\ \langle\vec{x}|\phi\rangle &= \langle 0|\Psi(\vec{x})|\phi\rangle \\ &= \phi(\vec{x}). \end{aligned} \tag{9.12}$$

Charge densities and currents can also be considered as operators. For particles of charge e,

$$\begin{aligned} \rho(\vec{x}) &= e\Psi^\dagger(\vec{x})\Psi(\vec{x}), \\ \vec{j}(\vec{x}) &= \frac{e\hbar}{2mi}\left((-\nabla\Psi^\dagger(\vec{x}))\Psi(\vec{x}) + \Psi^\dagger(\vec{x})\nabla\Psi(\vec{x})\right). \end{aligned} \tag{9.13}$$

If the state $|\phi\rangle$ and $|\chi\rangle$ are one-particle states, one can use the relations above to verify the previous relations,

$$\begin{aligned} \langle\chi|\rho(\vec{x})|\phi\rangle &= e\chi^*(\vec{x})\phi(\vec{x}), \\ \langle\chi|\vec{j}(\vec{x})|\phi\rangle &= \frac{e\hbar}{2mi}\left[\chi^*(\vec{x})\nabla\phi(\vec{x}) - (\nabla\chi^*(\vec{x}))\,\phi(\vec{x})\right]. \end{aligned} \tag{9.14}$$

9.3 Energy and the Hamiltonian

One can write the Hamiltonian for freely moving particles:

$$\begin{aligned} \boldsymbol{H_0} &= -\frac{\hbar^2}{2m}\int d^3x\Psi^\dagger(\vec{x})\nabla^2\Psi(\vec{x}). \\ &= \frac{\hbar^2}{2mV}\int d^3x \sum_{\vec{k},\vec{k}'} k^2 e^{i(\vec{k}-\vec{k}')\cdot\vec{x}} a^\dagger_{\vec{k}} a_{\vec{k}'}. \end{aligned} \tag{9.15}$$

The integral over $\vec{x}$ gives zero due to the varying phase unless $\vec{k} = \vec{k}'$, at which point the phase is unity and the integral over $\vec{x}$ cancels the volume in the denominator,

$$\boldsymbol{H_0} = \sum_{\vec{k}} \frac{\hbar^2 k^2}{2m} a^\dagger_{\vec{k}} a_{\vec{k}}. \tag{9.16}$$

Thus, even though the Hamiltonian, expressed in terms of the field operators Ψ and $\Psi^\dagger$, looks like the familiar expression for a wave function, it is far more powerful as it correctly expresses the energy even when many particles are present in the system, and even in those cases where multiple particles occupy the same single-particle level.

9.4 Interaction with an External Potential

An interaction with an external potential can be written as

$$\boldsymbol{H}_{\text{int}} = \int \boldsymbol{d}^3\boldsymbol{r}\boldsymbol{\mathcal{V}}(\vec{\boldsymbol{r}})\boldsymbol{\Psi}^\dagger(\vec{\boldsymbol{r}})\boldsymbol{\Psi}(\vec{\boldsymbol{r}}). \tag{9.17}$$

This can be written in terms of momentum-space creation and destruction operators by substituting the expressions for $\boldsymbol{\Psi}^\dagger$ from Eq. (9.6),

$$\begin{aligned}\boldsymbol{H}_{\text{int}} &= \frac{\boldsymbol{1}}{\boldsymbol{V}}\int \boldsymbol{d}^3\boldsymbol{r}\sum_{\vec{\boldsymbol{k}},\vec{\boldsymbol{k}}'}\boldsymbol{a}^\dagger_{\vec{\boldsymbol{k}}}\boldsymbol{a}_{\vec{\boldsymbol{k}}'}\boldsymbol{e}^{-\boldsymbol{i}(\vec{\boldsymbol{k}}-\vec{\boldsymbol{k}}')\cdot\vec{\boldsymbol{r}}}\boldsymbol{\mathcal{V}}(\vec{\boldsymbol{r}}) \\ &= \frac{\boldsymbol{1}}{\boldsymbol{V}}\sum_{\vec{\boldsymbol{k}},\vec{\boldsymbol{k}}'}\tilde{\boldsymbol{\mathcal{V}}}(\vec{\boldsymbol{k}}-\vec{\boldsymbol{k}}')\boldsymbol{a}^\dagger_{\vec{\boldsymbol{k}}}\boldsymbol{a}_{\vec{\boldsymbol{k}}'},\end{aligned} \tag{9.18}$$

where $\tilde{\boldsymbol{\mathcal{V}}}$ is the Fourier transform of $\boldsymbol{\mathcal{V}}$. Thus, the form of the interaction is similar to that of the Ex. 9.1 in that this features terms where one state is created and another destroyed.

In fact, one can reduce the solution of the problem (finding the energy eigenvalues) to that of diagonalizing a matrix.

$$\boldsymbol{H} = \sum_{\boldsymbol{i},\boldsymbol{j}}\boldsymbol{A}_{\boldsymbol{j}\boldsymbol{i}}\boldsymbol{a}^\dagger_{\boldsymbol{j}}\boldsymbol{a}_{\boldsymbol{i}}. \tag{9.19}$$

The simplified Hamiltonian results by linearly transforming the states to a new basis where $\boldsymbol{A} \to \boldsymbol{B}$, and $\boldsymbol{B}$ is diagonalized. In this new basis, one transforms the operators with a unitary transformation $\boldsymbol{U}$,

$$\begin{aligned}\boldsymbol{a}_{\boldsymbol{i}} &= \boldsymbol{U}_{\boldsymbol{i}\boldsymbol{i}'}\boldsymbol{b}_{\boldsymbol{i}'}, \\ \boldsymbol{a}^\dagger_{\boldsymbol{j}} &= \boldsymbol{U}^\dagger_{\boldsymbol{j}\boldsymbol{j}'}\boldsymbol{b}^\dagger_{\boldsymbol{j}'}, \\ \boldsymbol{H} &= \boldsymbol{b}^\dagger_{\boldsymbol{j}'}\boldsymbol{U}^\dagger_{\boldsymbol{j}'\boldsymbol{j}}\boldsymbol{A}_{\boldsymbol{j}\boldsymbol{i}}\boldsymbol{U}_{\boldsymbol{i}\boldsymbol{i}'}\boldsymbol{b}_{\boldsymbol{i}'}.\end{aligned} \tag{9.20}$$

If the transformation diagonalizes the Hamiltonian, where the diagonalized Hamiltonian is $\boldsymbol{B}_{\boldsymbol{\ell}\boldsymbol{\ell}'} = \boldsymbol{B}_{\boldsymbol{\ell}\boldsymbol{\ell}}\boldsymbol{\delta}_{\boldsymbol{\ell}\boldsymbol{\ell}'}$,

$$\begin{aligned}\boldsymbol{B}_{\boldsymbol{\ell}\boldsymbol{\ell}'} &= \boldsymbol{U}^\dagger_{\boldsymbol{\ell}\boldsymbol{j}}\boldsymbol{A}_{\boldsymbol{j}\boldsymbol{i}}\boldsymbol{U}_{\boldsymbol{i}\boldsymbol{i}\boldsymbol{\ell}'}, \\ \boldsymbol{H} &= \sum_{\boldsymbol{\ell}}\boldsymbol{B}_{\boldsymbol{\ell}\boldsymbol{\ell}}\boldsymbol{b}^\dagger_{\boldsymbol{\ell}}\boldsymbol{b}_{\boldsymbol{\ell}}.\end{aligned} \tag{9.21}$$

Thus, the same transformation used to diagonalize the matrix $\boldsymbol{A}_{\boldsymbol{i}\boldsymbol{j}}$ can be used to transform the operators to a basis where the Hamiltonian looks like a series of independent modes $\boldsymbol{\ell}$, where each mode can have any number of particles, $\boldsymbol{n}_{\boldsymbol{\ell}}$, with energies $\boldsymbol{E}_{\boldsymbol{\ell}}(\boldsymbol{n}_{\boldsymbol{\ell}}) = \boldsymbol{n}_{\boldsymbol{\ell}}\boldsymbol{B}_{\boldsymbol{\ell}\boldsymbol{\ell}}$.

One should keep in mind that even if there are only a few single-particle levels, one still has an infinite number of states because each mode $\boldsymbol{\ell}$ can hold an arbitrary number of particles. The ability of this formalism to include an arbitrary number of particles makes it the starting point for any study of many-body physics. It also represents the starting point for the study of relativistic physics due to the fact that even the vacuum might contain an arbitrary number of virtual particle-antiparticle pairs.

The true challenges for many-body physics arise when the interaction is not that with an external potential but when the potential has explicit interactions between constituents. The interaction term then has terms to higher powers in the field operators, e.g $\int d^3r d^3r'\ V(\vec{r}-\vec{r'})\psi^\dagger(\vec{r})\psi^\dagger(\vec{r'})\psi(\vec{r'})\psi(\vec{r})$. The problem cannot then be solved by rearranging the single-particle states into a basis where the Hamiltonian is diagonalized.

The most common example where particles are created or destroyed is in the radiative decay of atomic or nuclear states. For example, a hydrogen atom in an excited state might decay to the ground state via the emission of a photon. The photon must be created in the process. Creation of a photon is surprisingly complicated due to the fact that the photon is a massless spin-1 particle coupled to the current via the polarization vector. Because the algebraic gymnastics of polarization and spins adds additional complication, we will first consider the following example of a decay through the emission of a scalar (spinless) particle.

Example 9.2: Decays Via Emission of a Created Particle, Radiation

Consider a hypothetical massive particle, of mass m, created by the field operator $\Psi^\dagger$ and destroyed by Ψ. It is coupled to operators for a massless particle which is created by the field operator $\Phi^\dagger$ and destroyed by Φ. We assume both particles are spinless. The massive particle is in the first excited state of a three-dimensional Harmonic oscillator characterized by frequency ω_0. There first-excited state is three-fold degenerate, and in the Cartesian basis these states can be written as a product of three one-dimensional wave functions, $\psi_a(x)\psi_b(y)\psi_c(z)$. For this problem, we assume the excited state applies to the $z-$dependence, i.e. a and b reference the ground state and c refers to the first excited state. We wish to calculate the rate for decaying to the ground state via the emission of the massless particle. The interaction term is expressed as

$$H_{\text{int}} = g\int d^3r \Psi^\dagger(\vec{r})\left(\Phi^\dagger(\vec{r})+\Phi(\vec{r})\right)\Psi(\vec{r}),$$

which is manifestly Hermitian by noticing that $\Phi^\dagger(\vec{r})+\Phi(\vec{r})$ is Hermitian. The factor g is typically referred to as a coupling constant.
Using Fermi's golden rule the decay rate is

$$\Gamma = \frac{2\pi}{\hbar}\sum_{k_\beta}|\langle n_x, n_y, n_z=0; k_\beta|H_{\text{int}}|n_x, n_y=0, n_z=1\rangle|^2\,\delta(E_1-E_0-E_{k_\beta}),$$

where E_0 and E_1 are the energies of the ground state energy and of the first excited state, and $E_{k\beta}$ is the energy of the emitted particle. Because we are interested in the decay of the atom, which creates a particle through $\Phi^\dagger$, the destruction term, Φ, can be neglected. Also, note that the initial state was assumed to be polarized along the z axis. This choice is arbitrary because we are summing equally over all directions of k. If we were calculating a differential decay rate, $d\Gamma/d\Omega_k$, the angular dependence would depend on the initial polarization.
One can now calculate the matrix element,

$$\begin{aligned}&\langle n_x, n_y, n_z=0; k|H_{\text{int}}|n_x, n_y=0, n_z=1\rangle\\ &\quad= g\int d^3r\langle n_x, n_y, n_z=0; 0|a_k\Psi^\dagger(\vec{r})(\Phi^\dagger(\vec{r})+\Phi(\vec{r}))\Psi(\vec{r})|n_x, n_y=0, n_z=1\rangle,\end{aligned}$$

where the state $\langle \cdots ; \boldsymbol{k}|$ was re-expressed as $\langle \cdots ; 0|a_k$. Expanding the operator $\Phi^\dagger$ according to Eq. (9.6), and using Eq. (9.12) to bring out the wave functions, $\Psi(\vec{r})|\psi\rangle = \psi(\vec{r})|0\rangle$,

$$\begin{aligned}\langle n_x, n_y, n_z = 0; \boldsymbol{k}|H_{\text{int}}|n_x, n_y = 0, n_z = 1\rangle \\ = g \int d^3r \psi_0^*(\vec{r}) \frac{1}{\sqrt{V}} e^{-i\vec{k}\cdot\vec{r}} \psi_{n_z=1}(\vec{r}).\end{aligned}$$

Here, the harmonic oscillator wave functions can be written as a product of wave functions of x, y and z respectively. If the normalized 1-d harmonic oscillator wave functions are denoted $\psi_n(x)$,

$$\begin{aligned}&\langle n_x, n_y, n_z = 0; \boldsymbol{k}|H_{\text{int}}|n_x, n_y = 0, n_z = 1\rangle \\ &= \frac{g}{\sqrt{V}} \int dx\, e^{-ik_x x} \psi_0^2(x) \int dy\, e^{-ik_y y} \psi_0^2(y) \int dz\, e^{-ik_z z} z \sqrt{\frac{2m\omega_0}{\hbar}} \psi_0^2(z) \\ &= \frac{igk_z\alpha}{\sqrt{2V}} e^{-k^2\alpha^2/4},\end{aligned}$$

where α is the characteristic size of the ground state,

$$\alpha^2 \equiv \frac{\hbar}{m\omega_0}.$$

Putting all this together

$$\begin{aligned}\Gamma &= \frac{\pi g^2}{m\omega V} \sum_k k_z^2 e^{-k^2\alpha^2/2} \delta(\hbar kc - \hbar\omega_0) \\ &= \frac{g^2}{4m\omega\pi} \int k^4 dk \int_{-1}^{1} d\cos\theta\ \cos^2\theta\, e^{-k^2\alpha^2/2} \delta(\hbar kc - \hbar\omega) \\ &= \frac{g^2 k^3}{6\pi m\hbar c^2} e^{-k^2\alpha^2/2},\end{aligned}$$

where $k = \omega/c$, and ω is determined by energy conservation, $\hbar\omega = E_1 - E_0$.

9.5 Electromagnetic Decays

Electromagnetic decays are further complicated by the $\vec{j}\cdot\vec{A}$ nature of the coupling. Remember that minimal substitution, $\vec{p} \to (\vec{p} - e\vec{A}/c)$, results in a term in the Hamiltonian which looks like

$$H_{\text{int}} = -\int d^3r \left\{ \vec{j}(\vec{r})\cdot\vec{A}(\vec{r})/c + \frac{e^2}{2mc^2} \vec{A}(\vec{r})^2 \rho(\vec{r}) \right\}, \tag{9.22}$$

representing the interaction with the vector potential. The last term will be neglected for now, but plays a role in the quantum Hall effect. The current is given by the operator

$$\vec{j}(\vec{r}) = \frac{-ie\hbar}{2m}\left\{\Psi^\dagger(\vec{r})\nabla\Psi(\vec{r}) - [\nabla\Psi^\dagger(\vec{r})]\Psi(\vec{r})\right\}, \tag{9.23}$$

where $\Psi^\dagger$ and Ψ are field operators. Note that if sandwiched between the bra $\langle\psi|$ and the ket $|\psi\rangle$ one gets the usual expression for the current in terms of wave functions as $\Psi(\vec{r}) \to \psi(\vec{r})$ and $\Psi^\dagger(\vec{r}) \to \psi(\vec{r})$.

First we must define the electromagnetic field operator in terms of creation and destruction operators that make real photons,

$$\vec{A}(\vec{r},t) = \sqrt{\frac{2\pi\hbar^2c^2}{V}}\sum_{k,s}\frac{1}{\sqrt{E_k}}\left(\vec{\epsilon}_s(\vec{k})e^{i\vec{k}\cdot\vec{r}-iE_kt/\hbar}a_{k,s} + \vec{\epsilon}_s^*(\vec{k})e^{-i\vec{k}\cdot\vec{r}+iE_kt/\hbar}a_{k,s}^\dagger\right). \tag{9.24}$$

Here s refers to the polarization (or spin) of the photon. Polarization vectors can be complex. For example, for photons moving in the $\hat{z}$ direction, the polarization vector for right-circularly polarized light is $(\hat{x}+i\hat{y})/\sqrt{2}$. For each k there are two polarizations. Each must be perpendicular to the direction of $\vec{k}$, and normalized,

$$\vec{\epsilon}_s^*(\vec{k})\cdot\vec{\epsilon}_{s'}(\vec{k}) = \delta_{ss'}. \tag{9.25}$$

Aside from the polarization, the expression for $\vec{A}$ looks peculiar for the complicated prefactors and the $1/\sqrt{E_k}$ term inside the sum. However, all these factors are necessary to ensure that

$$\int d^3r\frac{\vec{E}^2+\vec{B}^2}{8\pi} = \sum_{\vec{k},s}E_k\left(a_{\vec{k},s}^\dagger a_{\vec{k},s} + \frac{1}{2}\right), \tag{9.26}$$

where the electromagnetic fields resulting from $\vec{A}$ are

$$\vec{E} = -\frac{1}{c}\frac{\partial\vec{A}}{\partial t},$$
$$\vec{B} = \nabla\times\vec{A}.$$

We are now in position to consider the general problem of electromagnetic decay, where a particle of charge e changes from state i to state f while emitting a photon of momentum k and polarization s. Outlining the steps to solving the problem:

1. Write down Fermi's golden rule,

$$\Gamma = \frac{2\pi}{\hbar}\sum_{\vec{k},s}|\frac{1}{c}\langle f;\vec{k},s|\int d^3r\,\vec{j}(\vec{r})\cdot\vec{A}(\vec{r})|i\rangle|^2\delta(E_i - E_f - \hbar kc). \tag{9.27}$$

2. The operator $\vec{A}$ operates on the photon degrees of freedom, while the current operator acts only on the charged particle's degrees of freedom. Thus, the matrix element can be written as a product,

$$\langle f;\vec{k},s|\frac{1}{c}\int d^3r\vec{j}(\vec{r})\cdot\vec{A}(\vec{r})|i\rangle = \int d^3r\,\langle f|\vec{J}(\vec{r})|i\rangle\cdot\langle\vec{k}|\vec{A}(\vec{r})/c|0\rangle. \tag{9.28}$$

Using the definition of the electromagnetic field operator in Eq. (9.24) one can easily write $\langle \vec{k}, s|\vec{A}(\vec{r})/c|0\rangle$ in terms of the plane wave function for the outgoing photon, and similarly, one can use write the current operator in terms of field operators to express $\langle f|\vec{J}(\vec{r})|i\rangle$ in terms of the initial and final wave functions for the massive particle,

$$\langle f|\vec{J}(\vec{r})|i\rangle = \frac{-ie\hbar}{2m}\left\{\psi_f^*(\vec{r})\nabla\psi_i(\vec{r}) - [\nabla\psi_f^*(\vec{r})]\psi_i(\vec{r})\right\}, \tag{9.29}$$

$$\langle \vec{k}, s|\vec{A}(\vec{r})/c|0\rangle = \sqrt{\frac{2\pi\hbar^2}{E_k V}}\vec{\epsilon}_s^*(\vec{k})e^{-i\vec{k}\cdot\vec{r}}.$$

The overall matrix element then becomes

$$\langle f;\vec{k},s|\frac{1}{c}\int d^3r\ \vec{j}(\vec{r})\cdot\vec{A}(\vec{r})|i\rangle = \frac{e\hbar}{m}\sqrt{\frac{2\pi}{E_k V}}\,\vec{\epsilon}_s^*(\vec{k})\cdot\vec{\mathcal{M}}. \tag{9.30}$$

$$\vec{\mathcal{M}}(\vec{k},i,f) \equiv \frac{\hbar}{2i}\int d^3r\ e^{-i\vec{k}\cdot\vec{r}}\psi_f^*(\vec{r})\nabla\psi_i(\vec{r}) - \frac{\hbar}{2i}\int d^3r\ e^{-i\vec{k}\cdot\vec{r}}[\nabla\psi_f^*(\vec{r})]\psi_i(\vec{r}).$$

Note $\mathcal{M}$ does not depend on the polarization, but does depend on the photon's momentum due to the factor $e^{-i\vec{k}\cdot\vec{r}}$.

3. Change the sum over $\vec{k}$ to an integral.

$$\sum_{\vec{k}} \to \frac{V}{(2\pi)^3}\int d^3k = \frac{V}{(2\pi)^3}\int k^2 dk d\Omega_k \tag{9.31}$$

Then eliminate the delta function in Fermi's golden rule by integrating over k,

$$\Gamma = \frac{2\pi}{\hbar}\frac{2\pi e^2\hbar^2}{V}\frac{V}{(2\pi)^3}\int k^2 dk\frac{1}{E_k}d\Omega_k\sum_s |\vec{\epsilon}_s^*(\vec{k})\cdot\vec{\mathcal{M}}|^2\delta(E_i - E_f - \hbar ck) \tag{9.32}$$

$$= \frac{e^2 k}{2\pi\hbar m^2c^2}\int d\Omega_k\sum_s |\vec{\epsilon}_s^*(\vec{k})\cdot\vec{\mathcal{M}}|^2.$$

One can check the units of the above expression by noting that e^2k has dimensions of energy (as it is the same dimension as e^2/r) and that $\mathcal{M}$ has units of momentum which are the same as the units of mc. Thus, the overall expression has units of energy over $\hbar$ which is an inverse time. Remember that in many books and tables Γ often refers to energies rather than rates, in which case one erases the $\hbar$ in the denominator.

4. To find the polarization for a specific direction, Ω, and for a specific polarization s, one can rewrite the expression without integrating over the solid angle or summing over polarization.

$$\frac{d\Gamma_s}{d\Omega_k} = \frac{e^2 k}{2\pi\hbar m^2c^2}\left|\vec{\epsilon}_s^*(\vec{k})\cdot\vec{M}\right|^2. \tag{9.33}$$

If one wishes to sum over the polarization directions, one can replace $\vec{M}$ with

$$\vec{\mathcal{P}} = \vec{M} - \hat{k}(\hat{k} \cdot \vec{\mathcal{M}}). \tag{9.34}$$

Because both polarization vectors are orthogonal to the direction of the photon, $\hat{k}$, this does not change the answer. The sum over s can now include the third polarization vector $\hat{k}$, and

$$\begin{aligned} \frac{d\Gamma}{d\Omega_k} &= \frac{e^2 k}{2\pi\hbar m^2 c^2} \sum_{s=1,2,3} |\vec{\epsilon}_s^* \cdot \vec{\mathcal{P}}|^2 \\ &= \frac{e^2 k}{2\pi\hbar m^2 c^2} |\vec{\mathcal{P}}|^2 \\ &= \frac{e^2 k}{2\pi\hbar m^2 c^2} \left\{ |\vec{\mathcal{M}}|^2 - |\hat{k} \cdot \vec{\mathcal{M}}|^2 \right\}. \end{aligned} \tag{9.35}$$

Even though there are two polarizations for any given photon direction, $\hat{k}$, the vector $\vec{\mathcal{P}}$ projects out a single direction. Further, because $\vec{\mathcal{P}}$ is orthogonal to $\hat{k}$, the polarization vector of the photon is indeed perpendicular to $\vec{k}$. Thus, knowing the directions of $\vec{\mathcal{M}}$ and $\hat{k}$ sets the polarization of the outgoing light to be

$$\vec{\epsilon}(\hat{k}) = \vec{\mathcal{P}}/|\vec{\mathcal{P}}|. \tag{9.36}$$

Note that if $\hat{k}$ is parallel to $\vec{\mathcal{M}}$, that $\vec{\mathcal{P}} = 0$ and the intensity vanishes in that direction.

5. Finally, to obtain the total rate Γ one must integrate over all directions of the photon. For any fixed vector $\vec{V}$, the integral of $|\hat{k} \cdot \vec{V}|^2$ over all directions of $\hat{k}$ is one third of the integral of $|\vec{V}|^2$, so the integrated decay rate is

$$\Gamma = \frac{4e^2 k}{3\hbar m^2 c^2} |\vec{\mathcal{M}}|^2. \tag{9.37}$$

9.6 The Dipole Approximation

For nuclear examples, typical γ energies are on the order of one MeV, which corresponds to wavelengths of a few hundred fm, approximately 100 times the size of a nucleus. For atomic examples, emitted photons usually have wavelengths of hundreds of nm, several thousand times the size of a typical atom. Thus, one might consider approximating the phase factor,

$$e^{-i\vec{k}\cdot\vec{r}} \approx 1. \tag{9.38}$$

This approximation is known as the dipole approximation,

$$\begin{aligned} \vec{\mathcal{M}} &\approx \langle f|\vec{P}|i\rangle \\ &= \frac{im}{\hbar}\langle f|[H_0, \vec{r}]|i\rangle \\ &= \frac{im(E_f - E_i)}{\hbar}\langle f|\vec{r}|i\rangle, \end{aligned} \tag{9.39}$$

thus resulting in a matrix element where $\vec{r}$ is sandwiched between the initial and final states, hence the term "dipole". The second line requires commuting $\boldsymbol{H_0 = \vec{P}^2/2m}$ with $\vec{r}$, using the fact that

$$\begin{aligned}
[|\vec{P}|^2, r_i] &= \sum_j [P_j^2, r_i] \\
&= \sum_j (P_j^2 r_i - r_i P_j^2) \\
&= \sum_j (P_j(r_i P_j + [P_j, r_i]) - r_i P_j^2) \\
&= \sum_j (P_j(r_i P_j - i\hbar\delta_{ij}) - r_i P_j^2) \\
&= -i\hbar P_i + \sum_j (r_i P_j^2 + [P_j, r_i]P_j - r_i P_j^2) \\
&= -2i\hbar P_i,
\end{aligned} \tag{9.40}$$

or $[\vec{P}^2, \vec{r}] = -2i\hbar\vec{P}$.

One consequence of dipole radiation is the dipole sum rule. Using the expression for the matrix element in Eq. (9.39),

$$|\langle f|\vec{r}|i\rangle|^2 = \frac{|\vec{\mathcal{M}}|^2\hbar^2}{m^2(E_f - E_i)^2}, \tag{9.41}$$

and rearranging Eq. (9.32),

$$\begin{aligned}
\int d\Omega_k \sum_s |\vec{\epsilon}_s^* \cdot \vec{\mathcal{M}}|^2 &= \Gamma_{i\to f}\frac{2\pi\hbar m^2c^2}{e^2k}, \\
&= \frac{8\pi}{3}|\vec{\mathcal{M}}|^2.
\end{aligned} \tag{9.42}$$

The last step introduced a factor of $2/3$ from averaging over the directions of $\vec{k}$, using the fact that $\vec{\epsilon}_s^*(\vec{k})$ is orthogonal to $\vec{k}$. Combining these last two expressions, and replacing k with $(E_i - E_f)/\hbar c$,

$$|\langle f|\vec{r}|i\rangle|^2 = \frac{3\hbar^4c^3}{4e^2(E_i - E_f)^3}\Gamma_{i\to f}. \tag{9.43}$$

Finally, if one sums over all states i, one can use completeness to reduce the l.h.s. to

$$e\langle f|r^2|f\rangle = \sum_i \frac{3\hbar^4c^3}{4e(E_i - E_f)^3}\Gamma_{i\to f}. \tag{9.44}$$

This is the dipole sum rule, and it relates the charge radius squared to a weighted sum over rates from all initial states i to some final state f. Although the derivation suggests that emission is through a single charge e, it can be derived more generally so that the l.h.s. is the charge radius

squared even for systems with charge spread over multiple particles with complicated wave functions. For objects composed of multiple charges, collective states can dominate the sum on the right hand side. When there are many charges that add to Z, and if those charges move together, i.e. collectively, they emit proportional to Z^2e^2, as opposed to single-particle motion which emits proportional to e^2. For example, a giant dipole resonance in ^{208}Pb nuclei involves the 82 protons oscillating together, while the 126 neutrons oscillate in the opposite direction. Such a state also absorbs photons readily.

Example 9.3: Electromagnetic Decay of the Harmonic Oscillator
Consider a harmonic oscillator in the first excited state, where in Cartesian coordinates the wave function is excited in the $\hat{z}$ direction. The wave function is then

$$\psi_1(\vec{r}) = \phi_0(x)\phi_0(y)\phi_1(z),$$

where ψ_0 and ψ_1 represent solutions to the one-dimensional harmonic oscillator.

1. Within the dipole approximation, calculate the rate per solid angle for decays of unpolarized photons.
Solution:
Using Eq.s (9.30) and (9.32) with $e^{-i\vec{k}\cdot\vec{r}} = 1$ by applying the dipole approximation, Eq. (9.39),

$$\begin{aligned}\Gamma &= \frac{e^2k}{2\pi\hbar m^2c^2}\int d\Omega_k \sum_s |\vec{\epsilon}_s^* \cdot \vec{\mathcal{M}}|^2,\\ \vec{\mathcal{M}}(\vec{k},i,f) &= \int d^3r\, \psi_f^*(\vec{r})\frac{\hbar}{i}\nabla\psi_i(\vec{r}),\\ &= im\omega\int d^3r\, \psi_f^*(\vec{r})\vec{r}\psi_i(\vec{r}).\end{aligned}$$

In the last step the change in energy of the two steps was replaced with the photon energy, $E_1 - E_0 = \hbar\omega$. Using the fact that the original state is excited along the $\hat{z}$ direction, one can then reduce the problem to a one-dimensional integral,

$$\vec{\mathcal{M}}(\vec{k},i,f) = im\omega\hat{z}\int dz\, \phi_0^*(z)z\phi_i(z).$$

Using solutions for the harmonic oscillator wave functions in Sec. (2.6) for the ground state and the excited states of the harmonic oscillator,

$$\vec{\mathcal{M}}(\vec{k},i,f) = i(\hbar m\omega/2)^{1/2}\hat{z}.$$

Inserting into the expression for Γ in Eq. (9.37),

$$\begin{aligned}\Gamma &= \frac{4e^2k}{3\hbar m^2c^2}\frac{\hbar m\omega}{2}\\ &= \frac{2e^2\omega^2}{3mc^3}.\end{aligned}$$

2. What is the angular distribution of photons? – summed over polarizations
Solution:
One can apply Eq. (9.35),

$$\begin{aligned}\frac{d\Gamma}{d\Omega} &= \frac{e^2 k}{2\pi\hbar m^2 c^2}\left\{|\vec{\mathcal{M}}|^2 - |\hat{k}\cdot\vec{\mathcal{M}}|^2\right\}, \\ &= \frac{e^2\omega^2}{4\pi m c^3}\left\{1-(\hat{k}\cdot\hat{z})^2\right\} \\ &= \frac{e^2\omega^2}{4\pi m c^3}\sin^2\theta_k.\end{aligned}$$

There are no photons emitted in the $\hat{z}$ direction because the matrix element $\vec{\mathcal{M}}$ is in the $\hat{z}$ direction and because the polarizations for a photon in the $\hat{z}$ direction must be perpendicular to $\hat{z}$ when $\hat{k}||\hat{z}$.

3. What is the polarization of a photon emitted in the $\hat{x}$ direction?
Solution:
By inspection, one can see that $\hat{\epsilon}$ must be parallel to $\hat{z}$. Thus, the photon would be plane polarized along the z direction.

9.7 Magnetic Dipole and Electric Quadrupole Radiation

The matrix element may be zero in the dipole approximation due to symmetry considerations, usually parity constraints. In that case, one should keep the next order term in the expansion of the factor $e^{-i\vec{k}\cdot\vec{r}} \approx 1 - i\vec{k}\cdot\vec{r} + \cdots$ in Eq. (9.30). One then obtains terms in the matrix element that look like

$$\begin{aligned}(\vec{k}\cdot\vec{r})(\vec{P}\cdot\vec{\epsilon}^*) &= \frac{1}{2}\left\{(\vec{k}\cdot\vec{r})(\vec{P}\cdot\vec{\epsilon}^*) - (\vec{\epsilon}^*\cdot\vec{r})(\vec{k}\cdot\vec{P})\right\} + \frac{1}{2}\left\{(\vec{k}\cdot\vec{r})(\vec{P}\cdot\vec{\epsilon}^*) + (\vec{\epsilon}^*\cdot\vec{r})(\vec{k}\cdot\vec{P})\right\} \\ &= \frac{1}{2}\left\{(\vec{k}\cdot\vec{r})(\vec{P}\cdot\vec{\epsilon}^*) - (\vec{\epsilon}^*\cdot\vec{r})(\vec{k}\cdot\vec{P})\right\} + \frac{1}{2}\left\{(\vec{P}\cdot\vec{\epsilon}^*)(\vec{k}\cdot\vec{r}) + (\vec{\epsilon}^*\cdot\vec{r})(\vec{k}\cdot\vec{P})\right\}.\end{aligned}$$

Here, the momentum operator $\vec{P} = -i\hbar\nabla$ acts on the operators of the charged particle, and should not be confused with $\hbar\vec{k}$, the momentum of the photon. The last step, which involved flipping the $(\vec{k}\cdot\vec{r})$ and $(\vec{P}\cdot\vec{\epsilon}^*)$ factors in the second bracket was justified by the fact that $\vec{P}$ commuting with $\vec{k}\cdot\vec{r}$ yields a result proportional to $\vec{k}$ which is in turn orthogonal to $\vec{\epsilon}$.
Using a vector identity the first term can be written

$$\frac{1}{2}\left\{(\vec{k}\cdot\vec{r})(\vec{P}\cdot\vec{\epsilon}^*) - (\vec{\epsilon}^*\cdot\vec{r})(\vec{k}\cdot\vec{P})\right\} = (\vec{r}\times\vec{P})\cdot(\vec{k}\times\vec{\epsilon}^*). \tag{9.45}$$

This operator looks like $\vec{L}\cdot(\vec{\epsilon}^*\times\vec{k})/2$, thus it looks like the magnetic dipole operator multiplied into $\vec{\epsilon}^*\times\vec{k}$. Remembering that $\vec{A}$ is defined by the direction of ϵ, and that $\vec{k}\times\vec{A}$ is generated by $\nabla\times\vec{A}$, one can think of this term as $\vec{L}\times\vec{B}$, which motivates the term magnetic dipole.

Using the identity

$$\vec{\boldsymbol{P}} = \frac{\boldsymbol{m}}{\boldsymbol{i\hbar}}[\vec{\boldsymbol{r}}, \boldsymbol{H_0}], \tag{9.46}$$

the second term can be written as

$$\begin{aligned} \frac{\boldsymbol{1}}{\boldsymbol{2}}\langle \boldsymbol{f}| \left\{ (\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{r}})(\vec{\boldsymbol{P}}\cdot\vec{\boldsymbol{\epsilon}}^*) + (\vec{\boldsymbol{\epsilon}}^*\cdot\vec{\boldsymbol{r}})(\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{P}}) \right\} |\boldsymbol{i}\rangle \\ &= \frac{\boldsymbol{m}}{\boldsymbol{i\hbar}}(\boldsymbol{E_i}-\boldsymbol{E_f})\langle \boldsymbol{f}|(\vec{\boldsymbol{\epsilon}}^*\cdot\vec{\boldsymbol{r}})\cdot(\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{r}})|\boldsymbol{i}\rangle \\ &= \frac{\boldsymbol{m}}{\boldsymbol{i\hbar}}(\boldsymbol{E_i}-\boldsymbol{E_f})\boldsymbol{\epsilon_i^* k_j Q_{ij}}, \\ \boldsymbol{Q_{ij}} &\equiv \langle \boldsymbol{f}|\boldsymbol{r_i r_j}|\boldsymbol{i}\rangle - \frac{\boldsymbol{1}}{\boldsymbol{3}}\boldsymbol{\delta_{ij}}\langle \boldsymbol{f}|\boldsymbol{r^2}|\boldsymbol{i}\rangle. \end{aligned} \tag{9.47}$$

The final term in the quadrupole operator $\boldsymbol{Q_{ij}}$ which is proportional to the delta function does not contribute because $\vec{\boldsymbol{\epsilon}}$ and $\vec{\boldsymbol{k}}$ are orthogonal. Radiation through this tThe term quadrupole comes from the fact that all such operators can be written in terms of $\boldsymbol{Y_{\ell=2,m}s}$. Both the magnetic dipole term and the electric quadrupole term are linear when expanding in $\vec{\boldsymbol{k}}$. They are smaller than the electric dipole piece only because $\boldsymbol{kr}$ is small, as they are of the same order in the coupling constant $\boldsymbol{e}$.

The symmetry constraints of the radiation pattern will be a major point of discussion later in the course when we discuss the Wigner-Eckart theorem.

9.8 Exercises

1. Consider two oscillator levels described by the creation operators, $\boldsymbol{a_1^\dagger}$ and $\boldsymbol{a_2^\dagger}$, where the Hamiltonian is

   $$\boldsymbol{H} = \boldsymbol{\epsilon_1 a_1^\dagger a_1} + \boldsymbol{\epsilon_2 a_2^\dagger a_2} + \boldsymbol{\beta}(\boldsymbol{a_1^\dagger a_2^\dagger} + \boldsymbol{a_1 a_2}).$$

   Consider the operators

   $$\begin{aligned} \boldsymbol{b_1^\dagger} &\equiv \boldsymbol{\cosh\eta\, a_1^\dagger} + \boldsymbol{\sinh\eta\, a_2}, \\ \boldsymbol{b_2^\dagger} &\equiv \boldsymbol{\cosh\eta\, a_2^\dagger} + \boldsymbol{\sinh\eta\, a_1}. \end{aligned}$$

   (a) Show that $\boldsymbol{b_i}$ and $\boldsymbol{b_i^\dagger}$ behave like creation/destruction operators.

   (b) Find the values of $\boldsymbol{\eta}$, $\boldsymbol{E_0}$, $\boldsymbol{E_1}$ and $\boldsymbol{E_2}$ that allow $\boldsymbol{H}$ to be written as

   $$\boldsymbol{H} = \boldsymbol{E_0} + \boldsymbol{E_1 b_1^\dagger b_1} + \boldsymbol{E_2 b_2^\dagger b_2}.$$

   This is known as a Bogoliubov transformation.

2. Consider $\boldsymbol{b}$-particles of mass $\boldsymbol{m}$ confined by a one-dimensional harmonic oscillator potential characterized by a frequency $\boldsymbol{\omega}$. The $\boldsymbol{b}$ particles interact with massless and spinless $\boldsymbol{a}$-particles through their respective field operators,

   $$\boldsymbol{H_{\text{int}}} = \boldsymbol{g}\int \boldsymbol{dx}\Psi^\dagger(\boldsymbol{x})\Phi(\boldsymbol{x})\Psi(\boldsymbol{x}),$$

where Φ and Ψ are the field operators for the a-particles and b-particles respectively. Assume the b particles are sufficiently heavy to ignore their recoil energy.

$$\Phi(x) = \frac{1}{\sqrt{L}} \sum_k \frac{1}{\sqrt{E_k}} \left(e^{-ikx} a_k^\dagger + e^{ikx} a_k \right)$$
$$\Psi^\dagger(x) = \frac{1}{\sqrt{L}} \sum_k e^{-ikx} b_k^\dagger,$$

(a) What is the dimension of g?

(b) What is the decay rate of a b particle in the first excited state.

3. Show that Eq. (9.26) is satisfied by using the electric and magnetic fields defined in Eq. (9.24). Note: You will ignore those cross terms involving rapid oscillations in time.

4. A proton in a nucleus decays from an excited state to its ground state by emitting a photon of momentum $\hbar\vec{k}$ and polarization $\vec{\epsilon}_s$. The matrix element describing the decay is

$$\langle 0, k, s|V|1\rangle = \beta \vec{\epsilon}_s \cdot \int d^3r \frac{e^{-i\vec{k}\cdot\vec{r}}}{\sqrt{V}} \left(\phi_0^*(\vec{r}) \nabla \phi_1(\vec{r}) - [\nabla \phi_0^*(\vec{r})] \phi_1(\vec{r}) \right).$$

The factor β absorbed all the various factors involved in defining the vector field in Eq. (9.24). Assume the ground and excited states are modeled with a three-dimensional harmonic oscillator of frequency ω. If the excited state is in the first level of a harmonic oscillator and has an angular momentum projection m, what is the angular distribution of the photons, $d\Gamma/d\Omega$, for each m. Assume that the wavelength of the photon is sufficiently long that the phase $e^{i\vec{k}\cdot\vec{r}} \approx 1$. Remember that the two polarizations of the photon must be perpendicular to $\vec{k}$. You need only calculate the angular shape of the distribution – ignore the prefactors.
Some help: The first excited state of the harmonic oscillator is three-fold degenerate. In the Cartesian basis these have the form $\sim x\phi_0, y\phi_0$ and $z\phi_0$, where ϕ_0 is the ground-state wave function. These can be mapped to three states that are eigenstates of angular momentum, $\ell = 1; m = 1, 0, -1$ as discussed in Chapter 4. The wave functions of states with $m = \pm 1$ have to have an angular dependence given by $Y_{1\pm1} \sim \sin\theta e^{\pm i\phi}$, whereas the wavefunction for $m = 0$ has to be proportional to $Y_{10} \sim \cos\theta$. Using the fact that $r\cos\theta = z$ and $r\sin\theta e^{\pm i\phi} = x \pm iy$, the $m = \pm 1$ wave functions are proportional to $x \pm iy$, whereas the $m = 0$ wave is proportional to z.

5. A spinless particle of mass M and charge e is in the first excited state of a three-dimensional harmonic oscillator characterized by a frequency ω. Assume the harmonic oscillator in the Cartesian state with $n_z = 1$, i.e. $m = 0$. Using the interaction

$$H_{\rm int} = \vec{j} \cdot \vec{A}/c,$$

(a) Calculate the decay rate of the charged particle into the ground state of the oscillator in the dipole approximation.

(b) Calculate $d\Gamma/d\Omega$ as a function of the emission angles of the photon, θ and ϕ.

(c) In terms of the unit vectors $\hat{\boldsymbol{k}}$, $\hat{\boldsymbol{\theta}}$ and $\hat{\phi}$, the two polarization vectors which are allowed for emission of a photon at an angle $\boldsymbol{\theta}, \phi$ are $\hat{\boldsymbol{\theta}}$ and $\hat{\phi}$. For each polarization vector above, calculate $\boldsymbol{d\Gamma_s/d\Omega}$, the probability of decaying via emission of a photon emited in the $\boldsymbol{\theta}, \phi$ direction with polarization $\boldsymbol{s}$.

6. Again consider a spinless particle of mass $\boldsymbol{M}$ and charge $\boldsymbol{e}$ n the first excited state of a three-dimensional harmonic oscillator characterized by a frequency $\boldsymbol{\omega}$. However, this time assume the charged particle is originally in a state with angular momentum projection $\boldsymbol{m = +1}$ along the $\boldsymbol{z}$ axis. Using the interaction

$$\boldsymbol{H_{\text{int}} = \vec{j} \cdot \vec{A}/c},$$

and applying the dipole approximation,

(a) Find the decay rate $\boldsymbol{\Gamma}$ of the first excited state.
(b) Find the differential decay rate $\boldsymbol{d\Gamma/d\Omega}$.
(c) Describe the polarization of a photon emitted in the $\hat{\boldsymbol{x}}$ direction.
(d) Describe the polarization vector of a particle emitted in the $\hat{\boldsymbol{z}}$ direction.

10 Advanced Topics in Angular Momentum

10.1 Isospin

The angular momentum algebra developed last semester has a broad range of applications, many of which have nothing to do with rotating systems but instead deal with symmetries in some other context. One such example is isospin. In nuclear and particle physics isospin mimics the spin mechanics used to describe spin $1/2$ particles. In quark language, the up quark is an $I = 1/2, I_3 = 1/2$ particle while the down quark is an $I = 1/2, I_3 = -1/2$ particle. Here, the total isospin is I, which is analogous to the total angular momentum J, and the isospin projection is I_3, which is analogous to the spin projection M. In nuclear physics the proton and neutron form an iso-doublet with the proton being the $I_3 = 1/2$ particle. The strong interaction is invariant to "rotations" in isospin space. In Chapter 4 we showed that rotations of angle θ about axis in the $\hat{n}$ direction were generated with the operator $e^{-i\vec{L}n\theta/\hbar}$. For spin-1/2 particles, the two components states were rotated by the operator $e^{-i\vec{\sigma}\cdot\hat{n}\theta/2}$. If one expresses the proton and neutron as two-component states, the transformation $e^{-i\vec{\sigma}\cdot\hat{n}\theta/2}$ can be thought of as a rotation in isospin. One can combine particles with a given isospin to form states with a specific net isospin in the same way that angular momentum are coupled. This means that the symmetry is analogous to the rotational symmetries discussed with angular momentum, and if the rotational symmetry in isospin is conserved (sometimes referred to as being a "good" symmetry) then the total isospin of the system is conserved. For instance, composite particles (those made of several quarks) have good isospin, which is determined by adding the isospin of several particles. Three up or down quarks can be coupled to either $I = 1/2$ states, the proton or neutron, or to an $I = 3/2$ state, the Δ baryon, in the same way three spin-half particles can be coupled to either $J = 3/2$ or $J = 1/2$ multiplets. The Δ baryon multiplet, which has $I = 3/2$, consists of four states with isospin projection $I_3 = -3/2, -1/2, 1/2$ and $3/2$. Because the down quark has electric charge $-1/3$ and the up quark has electric charge $+2/3$, the four Δs are $\Delta^-\Delta^0$, Δ^+ and Δ^{++}.

If one considers two nucleons (a nucleon is a proton or a neutron), and assume perfect isospin symmetry, one would assign the pair a total isospin of $I = 0$ or $I = 1$, with projections $I_3 = -1, 0, 1$ for $I = 1$. This is the same as listing all the ways to combine two spin $1/2$ particles, which combine into a triplet with $S = 1$ or a singlet with $S = 0$. The state with $I = 1$ and $I_3 = 1$ is made of two protons, and the state with $I = 1, I_3 = -1$ is made of two neutrons. The state with $I = 1, I_3 = 0$ is comprised of one proton and one neutron in a symmetric fashion. By symmetric, the two particle wave function would look like $\Psi(x_1, p; x_2, n) = \Psi(x_1, n; x_2, p)$. One would find this state by taking the two-proton state, then applying the lowering operator, which would result in a symmetric state. The $I = 0, I_3 = 0$ state would be anti-symmetric. Because of the symmetry, one would expect the three $I = 1$ states to behave similarly, while the $I = 0$ state might be different. Thus, the deuteron, which is made of a proton and a neutron has an $I = 0$ bound state, while the three $I = 1$ combinations have no bound states. Because the deutreron's isospin wave function is anti-symmetric, and because nucleons are spin-half fermions, the overall wave function must be anti-symmetric. This means that the deuteron has a symmetric spin wave function, and has $S = 1$. The three $I = 1$ states then have an anti-symmetric spin wave function, with $S = 0$. The pn combination thus has a bound state with $I = 0, S = 1$, and resonant behavior just above threshold in the $S = 0$ channel. This resonant

state is part of an $I = 1$ iso-triplet, i.e. the same $S = 0$ resonance appears in the pp and nn systems.

Isospin is an approximate symmetry. The symmetry is broken because the up and down quarks have slightly different masses, and because of the Coulomb interaction, which is obviously different for up quarks (charge 2/3) and down quarks (charge -1/3). Thus, the proton and neutron masses, 938.27 and 939.57 MeV$/c^2$ respectively, are not identical.

Example 10.1: The ρ meson
Consider the $\rho^{+,0,-}$ mesons (they have charge $+1, 0, -1$), which form an isotriplet ($I = 1$) and can decay into two pions. Pions, $\pi^{+,0,-}$, also form an iso-triplet. Find the decay branches for each of the three $\rho \to 2\pi$ decays.
Solution: Because isospin is conserved in the strong interaction, and because the ρ^0 has total isospin $I = 1$ and projection $I_3 = 0$, the two-pion state must also have the same total isospin and projection. First, look at the ρ^+ or ρ^-. The only branchings for each decay are

$$\rho^+ \to \pi^+\pi^0, \quad \rho^- \to \pi^-\pi^0.$$

However, the ρ^0 might go into one of two branches

$$\rho^0 \to \pi^+\pi^- \text{ or } \pi^0\pi^0.$$

To solve for the relative strengths of the two branches, find how to add two pion isospins to form an $I = 1, I_3 = 0$ state. Start by writing the $I = 2, I_3 = 2$ state

$$|I = 2, I_3 = 2\rangle = |\pi^+, \pi^+\rangle,$$

then use the lowering operator to find the $I = 2, I_3 = 1$ state. The lowering operator behaves the same as for angular momentum,

$$I_-|I, I_3\rangle = \sqrt{I(I+1) - I_3(I_3-1)}|I, I_3 - 1\rangle.$$

Thus,

$$|I = 2, I_3 = 1\rangle = \frac{1}{\sqrt{2}}|\pi^0, \pi^+\rangle + \frac{1}{\sqrt{2}}|\pi^+, \pi^0\rangle.$$

By orthogonality,

$$|I = 1, I_3 = 1\rangle = \frac{1}{\sqrt{2}}|\pi^+, \pi^0\rangle - \frac{1}{\sqrt{2}}|\pi^0, \pi^+\rangle$$

This combination corresponds to the decay of the ρ^+ meson. By applying the lowering operator,

$$\begin{aligned}\sqrt{1\cdot(1+1) - 1\cdot(1-1)}|I = 1, I_3 = 0\rangle &= \frac{\sqrt{1\cdot(1+1) + 0\cdot(0-1)}}{2}|\pi^-\pi^+\rangle \\ &\quad - \frac{\sqrt{1\cdot(1+1) + 0\cdot(0-1)}}{2}|\pi^+\pi^-\rangle \\ |I = 1, I_3 = 0\rangle &= \frac{1}{\sqrt{2}}|\pi^-\pi^+\rangle - \frac{1}{\sqrt{2}}|\pi^+\pi^-\rangle\end{aligned}$$

The $I = 1, I_3 = 0$ combination refers to the ρ^0 and has no $\pi^0\pi^0$ contribution. Thus, the two pions must both be charged if they couple to the ρ^0.
Also, the fact that the two terms appear identical, except for the switching of the π^+ and π^- tells us that the two final spatial (momentum) states of the pions must be anti-symmetric to cancel the anti-symmetry of the isospin wave function. This is accomplished by requiring the ρ to decay into a p wave. Note that the ρ^+ and ρ^- isospin wave functions were also anti-symmetric.

Angular momentum algebra applies to any system where a "rotation", e.g. $e^{i\sigma_i\phi/2}$, leaves the Lagrangian unchanged. Isospin is not an "exact" symmetry. Although the strong interaction conserves isospin, the electromagnetic interaction manifestly violates isospin. Thus, the ρ^0 does not decay with 100% probability into two charged pion, but has a very small probability to decay into $\pi_0\pi_0$ that is not even listed in the particle data book, http://pdg.lbl.gov).

Conservation of isospin plays a large role in nuclear physics, where the proton and neutron form the basis. A nucleus like ^{12}C which has $I_3 = 0$ might have states of total isospin, $I = 0, 1, 2, \cdots 6$. The nucleus ^{12}N has $I_3 = 1$ so it cannot have $I = 0$. The ground state of ^{12}C has $I = 0$ and has no analog in Boron or Nitrogen, but there does exist an excited state of ^{12}C with $I = 1$ which has very similar properties to the ground states of ^{12}N and ^{12}B.

Finally, we should mention that sometimes symmetries involve more than just the rotation of two states into one another. For instance, although the strange quark is much more massive than the up and down quarks, one can consider the three quarks to form a basis. If one ignored the quark masses, this basis has a greater symmetry as it involves rotations among three constituents, SU(3), rather than the SU(2) symmetry discussed before. In SU(3) the mechanics are more complicated compared to the SU(2) mechanics used to study angular momentum or isospin.

10.2 Combining Three Angular Momenta

When we studied two angular momenta, we described the "addition" as a change of basis where the state with labels $|j_j, j_1, m_1, m_2\rangle$ was written as a linear combination of states with labels $|j_1, j_2, J, M\rangle$. When adding three angular momenta the change of basis is

$$|j_1, j_2, j_3, m_1, m_2, m_3\rangle \leftrightarrow |(j_1, j_2), j_3, J_{12}, J, M\rangle. \tag{10.1}$$

Again, M is the projection of the total angular momentum J with values from $-J$ to J.

In order to couple the three angular momenta j_1 and j_2 were first coupled to J_{12} before j_3 and J_{12} were coupled to J. Thus, J_{12} survives as a quantum number, which is necessary as the original state had six labels, which requires six labels for the final state.

The change of basis can be described in terms of Clebsch-Gordan coefficients. First we describe

an intermediate state with j_1 and j_2 coupled to J_{12}.

$$|j_1, j_2, J_{12}, j_3, M_{12}, m_3\rangle = \sum_{m_1,m_2} |j_1, j_2, j_3, m_1, m_2, m_3\rangle\langle j_1, j_2, j_3, m_1, m_2, m_3| \quad (10.2)$$
$$|j_1, j_2, J_{12}, j_3, M_{12}, m_3\rangle$$
$$= \sum_{m_1,m_2} \langle j_1, m_1, j_2, m_2|J_{12}M_{12}\rangle|j_1, j_2, j_3, m_1, m_2, m_3\rangle.$$

The first line above is simply a statement of completeness. To write the second line one uses the fact that the labels j_3, m_3 in the matrix element $\langle j_1, j_2, j_3, m_1, m_2, m_3|j_1, j_2, J_{12}, j_3, M_{12}, m_3\rangle$ are unaffected by this change of basis, and can be neglected. The matrix element can then be replaced by a Clebsch-Gordan coefficient. One can now consider the states with total angular momentum J created by coupling J_{12} to j_3.

$$|(j_1, j_2), j_3, J_{12}, J, M\rangle = \sum_{M_{12},m_3} |j_1, j_2, j_3, J_{12}, M_{12}, m_3\rangle\langle j_1, j_2, j_3, J_{12}, M_{12}, m_3| \quad (10.3)$$
$$|(j_1, j_2), j_3, J_{12}, J, M\rangle$$
$$= \sum_{m_1,m_2,m_3,M_{12}} \langle J_{12}, M_{12}, j_3, m_3|JM\rangle\langle j_1, m_1, j_2, m_2|J_{12}M_{12}\rangle$$
$$|j_1, j_2, j_3, m_1, m_2, m_3\rangle.$$

The choice of coupling j_1 with j_2 to J_{12} was arbitrary, as one might have instead chosen to couple j_1 to j_3 or j_2 to j_3. States in the $|j_1, j_2, j_3, J, J_{12}, M\rangle$ basis can be expressed in terms of states in the $|j_1, j_2, j_3, J, J_{23}, M\rangle$ basis This change of basis which can be represented by Racah coefficients W.

$$\langle (j_1, j_2), j_3, J_{12}, J, M|j_1, (j_2, j_3), J_{23}, J, M\rangle \quad (10.4)$$
$$= \delta_{M,M'}\sqrt{(2J_{12}+1)(2J_{23}+1)}W(j_1, j_2, j_3, J; J_{12}, J_{23})$$
$$= \delta_{M,M'}(-1)^{j_1+j_2+j_3+J}\sqrt{(2J_{12}+1)(2J_{23}+1)}\begin{Bmatrix} j_1 & j_2 & J_{12} \\ j_3 & J & J_{23} \end{Bmatrix},$$

where the quantity in brackets is known as the Wigner $6-j$ symbol, which is simply another incarnation of the Racah coefficient. As a homework problem you will be asked to express the Racah coefficient W in terms of Clebsch-Gordan coefficients.

10.3 Notation: Conventions for Clebsch-Gordan Coefficients and Wigner 3j or 6j Symbols

Unfortunately, there is no overwhelmingly accepted coefficient for writing Clebsh-Gordan coefficients. In order to stress that they are simply matrix elements, we have chosen to use the notation,

$$\langle j_1, m_1, j_2, m_2|JM\rangle = \langle j_1, m_1, j_2, m_2|j_1, j_2, J, M\rangle. \quad (10.5)$$

The indices j_1 and j_2 are not affected by the change of basis, thus they are typically not repeated in the ket above. By convention, the states $|JM\rangle$ are defined in such a way that the Clebsch-Gordan coefficients are real. Thus, the labels in the bra and the ket can be reversed. Sometimes semicolons are inserted between labels, and sometimes the commas above are left out. It is good practice to choose labels carefully so that is clear which projections (ms) are associated with which total angular momenta (js).

$$\begin{aligned}\langle j_1, m_1, j_2, m_2 | j_3 m_3\rangle &= C^{j_3 m_3}_{j_1,m_1,m_2,m_2} \\ &= (-1)^{-j_1+j_2-m_3}\sqrt{2j_3+1}\begin{Bmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{Bmatrix}.\end{aligned} \tag{10.6}$$

These latter notations mainly appeal to those with masochistic tendencies, and are often seen in older literature or textbooks.

Aside from the convention that Clebsch-Gordan coefficients are real, there is also a choice in sign. For example, when generating Clebsch-Gordan coefficients by using ladder operators, one first considered the highest possible total angular momentum J and highest projection M, then used lowering operators to generate those with lower M but the same J. Then to find those with J lowered by one, the matrix element was found by the orthogonality constraint. But this next element could have arbitrary sign. Thus, an additional convention is applied

$$\langle j_1, m_1 = j_1, j_2, m_2 = J - j_1 | JJ\rangle > 0. \tag{10.7}$$

The coefficients have a number of symmetry properties:

$$\begin{aligned}\langle j_1 m_1 j_2 m_2 | JM\rangle &= (-1)^{j_1+j_2-J}\langle j_1, -m_1, j_2, -m_2 | J, -M\rangle \\ &= (-1)^{j_1+j_2-J}\langle j_2, m_2, j_1, m_1 | JM\rangle \\ &= (-1)^{j_1-m_1}\sqrt{\frac{2J+1}{2j_2+1}}\langle j_1, m_1, J, -M | j_2, -m_2\rangle \\ &= (-1)^{j_2+m_2}\sqrt{\frac{2J+1}{2j_1+1}}\langle J, -M, j_2, m_2 | j_1, -m_1\rangle \\ &= (-1)^{j_1-m_1}\sqrt{\frac{2J+1}{2j_2+1}}\langle J, M, j_1, -m_1 | j_2, m_2\rangle \\ &= (-1)^{j_2+m_2}\sqrt{\frac{2J+1}{2j_1+1}}\langle j_2, -m_2, J, M | j_1, m_1\rangle.\end{aligned} \tag{10.8}$$

10.4 Irreducible Tensor Operators

Often physics involves calculating transition elements of the form

$$\langle \ell', m' | T^k_q | \ell, m\rangle, \tag{10.9}$$

where the labels in the kets and bras denote the angular momentum of the initial and final states and T^k_q is an operator that is part of a set of $(2k+1)$ operators, $T^k_{-k}, T^k_{-k+1}\cdots T^k_k$ that

transform like spherical harmonics $\boldsymbol{Y_{k,q}}$ under rotation. Such operators are known as *irreducible tensor operators*. Examples are

1. $\boldsymbol{L_z}$ which transforms like $\boldsymbol{Y_{1,0}}$. Three operators are

$$\begin{aligned} \boldsymbol{T^1_{\pm 1}} &= \boldsymbol{-\mp (L_x \pm i L_y)/\sqrt{2}}, \\ \boldsymbol{T^1_0} &= \boldsymbol{L_z}. \end{aligned}$$

2. $\boldsymbol{x}$ which transforms as $\boldsymbol{Y_{1,-1} - Y_{1,1}}$. The operators are

$$\begin{aligned} \boldsymbol{T^1_{\mp 1}} &= \boldsymbol{\mp (x \pm iy)/\sqrt{2}}, \\ \boldsymbol{T^1_0} &= \boldsymbol{z}, \\ \boldsymbol{x} &= \boldsymbol{(T^1_{-1} - T^2_1)/\sqrt{2}} \end{aligned}$$

3. $\boldsymbol{r^2}$ which transforms as $\boldsymbol{Y_{0,0}}$, i.e $\boldsymbol{T^0_0 = r^2}$.
4. $\boldsymbol{P_x P_y}$ which transform as $\boldsymbol{Y_{2,2} - Y_{2,-2}}$. The set of operators is

$$\begin{aligned} \boldsymbol{T^2_{\pm 2}} &= \boldsymbol{(P_x + i P_y)^2}, \\ \boldsymbol{T^2_{\pm 1}} &= \boldsymbol{\pm 2 P_z (P_x \pm i P_y)}, \\ \boldsymbol{T^2_0} &= \boldsymbol{(2P_z^2 - P_x^2 - P_y^2)\sqrt{2}}, \\ \boldsymbol{P_x P_y} &= \boldsymbol{-i(T^2_2 - T^2_{-2})/4}. \end{aligned}$$

One can express any analytic function of $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ as a sum of irreducible tensor operators. But not all functions can be expressed in terms of operators with a single value of $\boldsymbol{k}$. For example, writing $\boldsymbol{x^2 + y^2 = (2/3)(x^2 + y^2 + z^2) - (2z^2 - x^2 - y^2)/3}$ requires both $\boldsymbol{T^{k=0}}$ and $\boldsymbol{T^{k=2}}$ pieces. For the examples above, one can scale any of the sets of $\boldsymbol{2k+1}$ operators by any scalar function, e.g. multiply by a constant or by $\boldsymbol{r^2}$, and it is still a set of irreducible tensor operators. However, all the terms in the set must be scaled similarly to preserve the way in which the rotate into one another. The term "irreducible tensor operator" refers to the fact that under rotations the operators mix only amongst irreducible subsets,

$$\boldsymbol{R(\vec{\alpha}) T^k_q R^{-1}(\vec{\alpha}) = D^k_{q,q'}(\vec{\alpha}) T^k_{q'}}. \tag{10.10}$$

Here, the matrices $\boldsymbol{D^k_{q,q'}}$ describe the linear transformation of the $\boldsymbol{2k+1}$ elements of $\boldsymbol{T^k}$. The rotations mix only the $\boldsymbol{2k+1}$ operators with the same $\boldsymbol{k}$ but different $\boldsymbol{q}$. Specifically, "irreducible" refers to the fact that rotations can mix in any of the different $\boldsymbol{q}$ components. For example, if one considered the set of six operators $\boldsymbol{p_i p_j}$ (six and not nine because the matrix is symmetric) one could express any of the six operators in terms of the five $\boldsymbol{\ell = 2}$ operators, plus one $\boldsymbol{\ell = 0}$ operator, $\boldsymbol{p^2 = \vec{p} \cdot \vec{p}}$. The full set of six operators would be reducible, because the single $\boldsymbol{\ell = 0}$ operator does not mix with the other five during rotations. But, the set of five that transform as the $\boldsymbol{Y^{\ell=2}_m}$s would indeed be an irreducible set.

As stated above the form of the spherical harmonics can be used to express polynomials of $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ in terms of irreducible tensor operators. To that end, a list of the spherical harmonics and

associated polynomials to second order is presented here.

$$\begin{aligned}
\boldsymbol{Y_{0,0}} &= \frac{\mathbf{1}}{\sqrt{\mathbf{4\pi}}}, \\
\boldsymbol{Y_{1,0}} &= \sqrt{\frac{\mathbf{3}}{\mathbf{4\pi}}}\cos\boldsymbol{\theta}, \\
\boldsymbol{Y_{1,\pm 1}} &= \mp\sqrt{\frac{\mathbf{3}}{\mathbf{8\pi}}}\sin\boldsymbol{\theta} e^{\pm i\boldsymbol{\phi}}, \\
\boldsymbol{Y_{2,0}} &= \sqrt{\frac{\mathbf{5}}{\mathbf{16\pi}}}(\mathbf{3}\cos^{\mathbf{2}}\boldsymbol{\theta} - \mathbf{1}), \\
\boldsymbol{Y_{2,\pm 1}} &= \mp\sqrt{\frac{\mathbf{15}}{\mathbf{8\pi}}}\sin\boldsymbol{\theta}\cos\boldsymbol{\theta} e^{\pm i\boldsymbol{\phi}}, \\
\boldsymbol{Y_{2,\pm 2}} &= \sqrt{\frac{\mathbf{15}}{\mathbf{32\pi}}}\sin^{\mathbf{2}}\boldsymbol{\theta} e^{\pm \mathbf{2}i\boldsymbol{\phi}}, \\
\boldsymbol{Y_{\ell -m}}(\boldsymbol{\theta},\boldsymbol{\phi}) &= (-\mathbf{1})^{\boldsymbol{m}}\boldsymbol{Y^*_{\ell m}}(\boldsymbol{\theta},\boldsymbol{\phi}).
\end{aligned} \tag{10.11}$$

For $\boldsymbol{k} = \mathbf{0}, \mathbf{1}, \mathbf{2}$ one can define a set of irreducible tensor operators using powers of $\boldsymbol{x}, \boldsymbol{y}$ and $\boldsymbol{z}$. Defining

$$\boldsymbol{T^k_q} \equiv \boldsymbol{r^k Y_{k,q}}\sqrt{\mathbf{4\pi}/(\mathbf{2k}+\mathbf{1})}, \tag{10.12}$$

allows one to define the following polynomials in terms of irreducible tensor operators, using $\boldsymbol{z} = \boldsymbol{r}\cos\boldsymbol{\theta}$ and $(\boldsymbol{x} + \boldsymbol{iy}) = \boldsymbol{r}\sin\boldsymbol{\theta} e^{i\boldsymbol{\phi}}$.

$$\begin{aligned}
\boldsymbol{T^0_0} &= \mathbf{1}, \\
\boldsymbol{T^1_1} &= -\frac{\mathbf{1}}{\sqrt{\mathbf{2}}}(\boldsymbol{x} + \boldsymbol{iy}), \\
\boldsymbol{T^1_0} &= \boldsymbol{z}, \\
\boldsymbol{T^1_{-1}} &= \frac{\mathbf{1}}{\sqrt{\mathbf{2}}}(\boldsymbol{x} - \boldsymbol{iy}), \\
\boldsymbol{T^2_2} &= \sqrt{\frac{\mathbf{3}}{\mathbf{8}}}(\boldsymbol{x}^{\mathbf{2}} + \mathbf{2}\boldsymbol{ixy} - \boldsymbol{y}^{\mathbf{2}}), \\
\boldsymbol{T^2_1} &= -\frac{\sqrt{\mathbf{3}}}{\mathbf{2}}\boldsymbol{z}(\boldsymbol{x} + \boldsymbol{iy}), \\
\boldsymbol{T^2_0} &= \frac{\mathbf{1}}{\mathbf{2}}(\mathbf{3}\boldsymbol{z}^{\mathbf{2}} - \boldsymbol{r}^{\mathbf{2}}), \\
\boldsymbol{T^2_{-1}} &= \frac{\sqrt{\mathbf{3}}}{\mathbf{2}}\boldsymbol{z}(\boldsymbol{x} - \boldsymbol{iy}), \\
\boldsymbol{T^2_{-2}} &= \sqrt{\frac{\mathbf{3}}{\mathbf{8}}}(\boldsymbol{x}^{\mathbf{2}} - \mathbf{2}\boldsymbol{ixy} - \boldsymbol{y}^{\mathbf{2}}).
\end{aligned} \tag{10.13}$$

The prefactor $\boldsymbol{r}\sqrt{\mathbf{4\pi}(\mathbf{2k}+\mathbf{1})}$ was arbitrary. Any function of $\boldsymbol{k}$ is fine. It is only important that the ratio of elements within a given $\boldsymbol{k}$ multiplet are unchanged. These expressions can then be

inverted to express the polynomials in terms of irreducible tensor operators,

$$
\begin{aligned}
\mathbf{1} &= \boldsymbol{T}^{\mathbf{0}}_{\mathbf{0}}, \\
\boldsymbol{x} &= \frac{\mathbf{1}}{\sqrt{\mathbf{2}}}(\boldsymbol{T}^{\mathbf{1}}_{\mathbf{-1}} - \boldsymbol{T}^{\mathbf{1}}_{\mathbf{1}}), \\
\boldsymbol{y} &= \frac{\boldsymbol{i}}{\sqrt{\mathbf{2}}}(\boldsymbol{T}^{\mathbf{1}}_{\mathbf{-1}} + \boldsymbol{T}^{\mathbf{1}}_{\mathbf{1}}), \\
\boldsymbol{z} &= \boldsymbol{T}^{\mathbf{1}}_{\mathbf{0}}, \\
\boldsymbol{x^2} &= \frac{\mathbf{1}}{\mathbf{2}}\sqrt{\frac{\mathbf{2}}{\mathbf{3}}}(\boldsymbol{T}^{\mathbf{2}}_{\mathbf{2}} + \boldsymbol{T}^{\mathbf{2}}_{\mathbf{-2}}) - \frac{\mathbf{1}}{\mathbf{3}}\boldsymbol{T}^{\mathbf{2}}_{\mathbf{0}} + \frac{\mathbf{1}}{\mathbf{3}}\boldsymbol{T}^{\mathbf{0}}_{\mathbf{0}}\boldsymbol{r^2}, \\
\boldsymbol{y^2} &= -\frac{\mathbf{1}}{\mathbf{2}}\sqrt{\frac{\mathbf{2}}{\mathbf{3}}}(\boldsymbol{T}^{\mathbf{2}}_{\mathbf{2}} + \boldsymbol{T}^{\mathbf{2}}_{\mathbf{-2}}) - \frac{\mathbf{1}}{\mathbf{3}}\boldsymbol{T}^{\mathbf{2}}_{\mathbf{0}} + \frac{\mathbf{1}}{\mathbf{3}}\boldsymbol{T}^{\mathbf{0}}_{\mathbf{0}}\boldsymbol{r^2}, \\
\boldsymbol{z^2} &= \frac{\mathbf{2}}{\mathbf{3}}\boldsymbol{T}^{\mathbf{2}}_{\mathbf{0}} + \frac{\mathbf{1}}{\mathbf{3}}\boldsymbol{T}^{\mathbf{0}}_{\mathbf{0}}\boldsymbol{r^2}, \\
\boldsymbol{xy} &= \boldsymbol{i}\frac{\mathbf{1}}{\sqrt{\mathbf{6}}}(\boldsymbol{T}^{\mathbf{2}}_{\mathbf{-2}} - \boldsymbol{T}^{\mathbf{2}}_{\mathbf{2}}), \\
\boldsymbol{xz} &= \frac{\mathbf{1}}{\sqrt{\mathbf{3}}}(\boldsymbol{T}^{\mathbf{2}}_{\mathbf{-1}} - \boldsymbol{T}^{\mathbf{2}}_{\mathbf{1}}), \\
\boldsymbol{yz} &= \frac{\boldsymbol{i}}{\sqrt{\mathbf{3}}}(\boldsymbol{T}^{\mathbf{2}}_{\mathbf{-1}} + \boldsymbol{T}^{\mathbf{2}}_{\mathbf{1}}).
\end{aligned}
\tag{10.14}
$$

This procedure could have been used to express other polynomials, e.g. $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \rightarrow \boldsymbol{L_x}, \boldsymbol{L_y}, \boldsymbol{L_z}$, because they behave identically under rotation.

For operators created by taking products of two different operators, e.g. $\boldsymbol{P_x L_z}$, one needs be more careful, because in this case $\boldsymbol{P_x L_z}$ does not behave like $\boldsymbol{xz}$ under rotation. If one considers $\boldsymbol{P_i L_j}$ as a $\mathbf{3} \times \mathbf{3}$ matrix, there are 9 independent terms. In contrast the combination $\boldsymbol{P_i P_j}$ would have only 6 terms because it is manifestly symmetric. For rank-2 tensors combined from two different vector operators $\vec{\boldsymbol{A}}$ and $\vec{\boldsymbol{B}}$, one can consider the following combinations, $\vec{\boldsymbol{A}} \cdot \vec{\boldsymbol{B}}$, $\boldsymbol{A_i B_j} - \boldsymbol{B_i A_j}$ and $\boldsymbol{A_i B_j} + \boldsymbol{B_i A_j} - \mathbf{2}\vec{\boldsymbol{A}} \cdot \vec{\boldsymbol{B}}/\mathbf{3}$. The first combination, $\vec{\boldsymbol{A}} \cdot \vec{\boldsymbol{B}}$ is clearly a scalar, and behaves like an irreducible tensor operator, $\boldsymbol{T}^{\boldsymbol{k=0}}_{\boldsymbol{q=0}}$. The second is antisymmetric and behaves as part of the three pseudo-vector components, $\boldsymbol{\epsilon_{ijk} A_j B_k}$, or equivalently $\vec{\boldsymbol{A}} \times \vec{\boldsymbol{B}}$. These can be expressed as a linear sum of the three $\boldsymbol{T^{k=1}q}$ operators as was done above for $\boldsymbol{x}, \boldsymbol{y}$ and $\boldsymbol{z}$. This leaves the traceless symmetric combination at the end. These five independent operators can be written in terms of $\boldsymbol{T}^{\boldsymbol{k=2}}_{\boldsymbol{q}}$ operators. For example, the operator $\boldsymbol{P_x L_y}$ would be needed to be expressed as a sum over some pieces with some operators $\boldsymbol{T}^{\boldsymbol{k=2}}_{\boldsymbol{q}}$ and combined with contributions from a set of irreducible operators with $\boldsymbol{T'}^{\boldsymbol{k=1}}_{\boldsymbol{q}}$.

Example 10.2: Combining operators to make irreducible tensor operators
Express the operator $\boldsymbol{P_x L_y}$ as a linear sum over irreducible tensor operators. Following the discussion above

$$
\boldsymbol{P_i L_j} = \frac{\mathbf{1}}{\mathbf{3}}\vec{\boldsymbol{P}} \cdot \vec{\boldsymbol{L}}\boldsymbol{\delta_{ij}} + \left(\frac{\boldsymbol{P_i L_j} + \boldsymbol{P_j L_i}}{\mathbf{2}} - \frac{\mathbf{1}}{\mathbf{3}}\vec{\boldsymbol{P}} \cdot \vec{\boldsymbol{L}}\boldsymbol{\delta_{ij}}\right) + \left(\frac{\boldsymbol{P_i L_j} - \boldsymbol{P_j L_i}}{\mathbf{2}}\right).
$$

The first term is a scalar and is invariant under rotation, i.e. it does not mix with any other

operators. The product $\boldsymbol{P_x L_y}$ has contribution from this element. The second term will transform amongst a set of $\boldsymbol{k=2}$ irreducible tensor operators, while the third transforms like the components of a vector $\vec{\boldsymbol{A}} = \vec{\boldsymbol{P}} \times \vec{\boldsymbol{L}}/2$, and can be expressed as a set of $\boldsymbol{k=1}$ operators. Defining the $\boldsymbol{k=2}$ operators using Eq. (10.14) with both $\vec{\boldsymbol{P}}$ and $\vec{\boldsymbol{L}}$ being replaced by the components $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$,

$$(\boldsymbol{P_x L_y + P_y L_x})/2 = i\frac{1}{2\sqrt{6}}(\boldsymbol{S^2_{-2} - S^2_2}).$$

Here, $\boldsymbol{S^k_q}$ is defined as irreducible tensor operators using Eq. (10.13) where each term $\boldsymbol{r_i r_j}$ is replaced by $(\boldsymbol{P_i L_j + P_j L_i})/\boldsymbol{2}$. The third term is also expressed using Eq. (10.13),

$$\begin{aligned} \vec{\boldsymbol{A}} &\equiv \vec{\boldsymbol{P}} \times \vec{\boldsymbol{L}}/2, \\ \boldsymbol{A_z} &= \boldsymbol{Q^{k=1}_{q=0}}, \\ \pm(\boldsymbol{A_x} \pm i\boldsymbol{A_y})/\sqrt{2} &= \boldsymbol{Q^k_{\pm 1}}, \end{aligned}$$

and

$$(\boldsymbol{P_x L_y - P_y L_x})/2 = \boldsymbol{A_z} = \boldsymbol{Q^{k=1}_0}.$$

This illustrates the meaning of the term "irreducible". Under rotation, the nine operators $\boldsymbol{P_i L_j}$ rotate amongst each other, but the nine operators can be divided, or reduced, into the three sets, $\boldsymbol{k = 0, 1, 2}$, where an element of each set does not mix with an element of the other two sets under rotation.

10.5 Wigner $\mathcal{D}$ matrices

The rotation matrices $\boldsymbol{\mathcal{D}^k_{q,q'}}(\vec{\boldsymbol{\alpha}})$ can be defined in terms of Euler angles,

$$\begin{aligned} \boldsymbol{R}(\boldsymbol{\alpha_1, \alpha_2, \alpha_3}) &= \boldsymbol{e^{-i\alpha_1 J_z/\hbar} e^{-i\alpha_2 J_y/\hbar} e^{-i\alpha_3 J_z/\hbar}}, \\ \boldsymbol{\mathcal{D}^\ell_{m,m'}}(\vec{\boldsymbol{\alpha}}) &= \langle \boldsymbol{\ell m'} | \boldsymbol{R}(\vec{\boldsymbol{\alpha}}) | \boldsymbol{\ell, m} \rangle. \end{aligned} \tag{10.15}$$

The order of the rotations might seem odd above. The first Euler rotation is described by $\boldsymbol{\alpha_1}$, yet it appears to operate last on the ket. To understand why the Euler angles are expressed in the particular order, one can write

$$\boldsymbol{R}(\boldsymbol{\phi, \theta, \psi}) = \boldsymbol{e^{-iJ_{z'}\psi/\hbar} e^{-iJ_{y'}\theta/\hbar} e^{-iJ_z\phi/\hbar}}, \tag{10.16}$$

where the $\boldsymbol{y'}$ axis is the $\boldsymbol{y}$ axis after an initial rotation around the original $\boldsymbol{z}$ axis by an angle $\boldsymbol{\phi}$ and $\boldsymbol{z'}$ is the new $\boldsymbol{z}$ axis formed after the rotation of $\boldsymbol{\theta}$ about the $\boldsymbol{y'}$ axis. Thus,

$$\begin{aligned} \boldsymbol{e^{-iJ_{y'}\theta/\hbar}} &= \boldsymbol{e^{-iJ_z\phi/\hbar} e^{-iJ_y\theta/\hbar} e^{iJ_z\phi/\hbar}} \\ \boldsymbol{e^{-iJ_{z'}\psi/\hbar}} &= \boldsymbol{e^{-iJ_{y'}\theta/\hbar} e^{-iJ_z\psi/\hbar} e^{iJ_{y'}\theta/\hbar}}. \end{aligned} \tag{10.17}$$

Substituting these into the expression for $\boldsymbol{R}$ allows one to write the rotation matrices without mentioning primed axes. Aside from the primes disappearing, the angles $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$ have

swapped their order,

$$\boldsymbol{R}(\phi,\theta,\psi) = e^{-iJ_z\phi/\hbar}e^{-iJ_y\theta/\hbar}e^{-iJ_z\psi/\hbar}. \tag{10.18}$$

The $\mathcal{D}$ matrices, or Wigner $\mathcal{D}$ matrices, can be written in closed form,

$$\begin{aligned} D^j_{m'm}(\alpha_1,\alpha_2,\alpha_3) &= e^{-im'\alpha_1}d^j_{m'm}(\alpha_2)e^{-im\alpha_3}, \\ d^j_{m'm}(\alpha_2) &= (-1)^\lambda \begin{pmatrix} 2j-k \\ k+a \end{pmatrix}^{1/2} \begin{pmatrix} k+b \\ b \end{pmatrix}^{-1/2} \\ &\quad \cdot (\sin(\alpha_2/2))^a\,(\cos(\alpha_2/2))^b\,P_k^{(a,b)}(\cos\alpha_2), \end{aligned} \tag{10.19}$$

where

$$\begin{aligned} k &= \min(j+m, j-m, j+m', j-m'), \\ k &= \begin{cases} j+m: & a=m'-m; \quad \lambda = m'-m \\ j-m: & a=m-m'; \quad \lambda = 0 \\ j+m': & a=m-m'; \quad \lambda = 0 \\ j-m': & a=m'-m; \quad \lambda = m'-m \end{cases}, \\ b &= 2j-2k-a. \end{aligned}$$

Here, $P_k^{a,b}(\cos\alpha_2)$ are Jacobi polynomials. A description of their properties can be found in https://en.wikipedia.org/wiki/Jacobi_polynomials. Fortunately, routines for calculating $\mathcal{D}$ matrices are readily available. For more properties of $\mathcal{D}$ matrices one can visit https://en.wikipedia.org/wiki/Wigner_D-matrix.

The $\mathcal{D}$ matrices have a variety of properties, many of which are related to those of the spherical harmonics. For example, the orthogonality property is

$$\int d\omega\, \mathcal{D}^{(j_1)*}_{m_1',m_1}(\vec{\omega})\mathcal{D}^{(j_2)}_{m_2',m_2}(\vec{\omega}) = \frac{\delta_{j_1,j_2}\delta_{m_1,m_2}\delta_{m_1',m_2'}}{2j_1+1}. \tag{10.20}$$

Here the integral over $d\omega$ is shorthand for averaging all three Euler angles, i.e. the integration covers all the $0 < \alpha_1, \alpha_3 < 2\pi$, and $-1 < \cos\alpha_2 < 1$, then divides by the net angle, $8\pi^2$ so that $\int d\omega = 1$. Note that for half integer spins, the integrals must subtend twice the angular range.

10.6 The Two-Dimensional Analogy of the Wigner Eckart Theorem

Before launching into the Wigner Eckart theorem, it is much easier to consider the two-dimensional example. Consider bras and kets, which transform as $e^{im\phi}$, and additionally, consider operators, T_q, which also transform as $e^{iq\phi}$. For example the operator x can be written as a sum of two pieces, each of which transforms in this manner.

$$x = \frac{r}{2}(e^{i\phi} + e^{-i\phi}). \tag{10.21}$$

The two terms will be considered separately, the one that transforms as $e^{i\phi}$ and the one that transforms as $e^{-i\phi}$. In the three-dimensional analog, each term will be called an irreducible

tensor operator. Under rotations, the states transform as

$$\begin{aligned} \mathcal{R}(\phi)|m\rangle &= e^{im\phi}|m\rangle, \\ \langle m|\mathcal{R}^{-1}(\phi) &= \langle m|e^{-im\phi}, \end{aligned} \tag{10.22}$$

and the operators transform as

$$\mathcal{R}(\phi)T_q\mathcal{R}^{-1}(\phi) = e^{iq\phi}T_q. \tag{10.23}$$

Thus, under rotation,

$$\begin{aligned} \langle m'|T_q|m\rangle &\to \langle m'|\mathcal{R}^{-1}(\phi)\mathcal{R}(\phi)T_q\mathcal{R}^{-1}(\phi)\mathcal{R}(\phi)|m\rangle \\ &= e^{i(m+q-m')\phi}\langle m'|T_q|m\rangle. \end{aligned} \tag{10.24}$$

The expectation must not depend on the rotation angle ϕ. This was certainly expected, because if you rotate both the state and the operator by the same angle, nothing should change. Thus, the matrix element must be zero unless $m + q - m' = 0$. One could state this fact by

$$\langle \beta', m'|T_q|\beta, m\rangle \delta_{m',m+q} F(m, q), \tag{10.25}$$

where finding $F(m, q)$ would typically involve solving for the matrix element for each m and each q. The labels β and β' simply encapsulate all other labels needed to describe the states, but which are not affected by rotation, e.g. the radial wave function.

The proofs of the two-dimensional and three-dimensional versions have much in common, but the three-dimensional version is more powerful. Jumping ahead, the three-dimensional version is

$$\langle \tilde{\beta}, J, M|T^k_q|\beta, \ell, m_\ell\rangle = \langle k, q, \ell, m|JM\rangle F(\tilde{\beta}, \tilde{J}, k, \beta, J), \tag{10.26}$$

where F might need to be calculated for each of the five arguments. Like the two-dimensional version, there is a constraint on summing the projections, $m + q = M$, and is enforced by the Clebsch-Gordan coefficient. Unlike the three-dimensional version, the function F does not depend on the projections, m, M and q. This has to do with the fact that in three dimensions one can rotate the states and operator about the x or y axis, which results in mixing various values of m, but does not mix different values of J. Thus, one can calculate F for one combination of m, q and M, as long as the Clebsch-Gordan coefficient is non-zero, then generate all other combinations by taking ratios of Clebsch-Gordan coefficients. This proof is detailed below.

Example 10.3: Expressing a Two-Dimensional Function in Terms of Irreducible Tensor Operators

Express the function, $f(x, y) = xy$, in terms of irreducible tensor operators. I.e., in polar coordinates where the angular dependences are of the form $e^{i\phi}$.

Solution: First, write

$$\begin{aligned} x &= \frac{r}{2}(e^{i\phi} + e^{-i\phi}), \\ y &= \frac{r}{2i}(e^{i\phi} - e^{-i\phi}), \\ xy &= \frac{r^2}{4i}\left(e^{2i\phi} - e^{-2i\phi}\right). \end{aligned}$$

Thus, xy can be considered as two terms, one which rotates as $e^{2i\phi}$ and one which rotates as $e^{-2i\phi}$. The two terms would be considered separately. If one were to calculate the matrix element $\langle \beta', m'|xy|\beta, m\rangle$, the result would be zero unless m and m' differed by ± 2. If the operator in question had been $P_x P_y$, the conclusion would have been the same because it can also be written in terms of irreducible tensor operators with $q = \pm 2$.

We should qualify the statement above that the expectation of the matrix element $\langle \alpha|\mathcal{O}|\beta\rangle$ should not change under rotations if you rotate both the states and the operator, i.e.,

$$\langle \alpha|R^{-1}(\phi)\left[R(\phi)\mathcal{O}R^{-1}(\phi)\right]R(\phi)|\beta\rangle = \langle \alpha|\mathcal{O}|\beta\rangle. \tag{10.27}$$

Sometimes there is an external potential or field. The operator might be defined in terms of the field, e.g. the interaction with a magnetic field, $\alpha\vec{B}\cdot\vec{S}$. For the statement above to be true, one must also rotate the external field. If one factors out the magnetic field and only considers the operators that operate on the state information represented by the bra and ket, then one should only apply the Wigner-Eckart theorem to that portion of the matrix element. In this case $\langle \alpha|\vec{S}|\beta\rangle$. One would not treat $\vec{B}\cdot\vec{S}$ as being a scalar operator unless the source of the magnetic field was also being described by the bra and ket. Thus, one separates the "external" operators from the "internal" ones before applying either the two- or three-dimensional Wigner-Eckart theorems.

10.7 The Wigner Eckart Theorem

The Wigner-Eckart theorem shows that the matrix element,

$$\langle \tilde{\beta}, J, M|T^k_q|\beta, \ell, m_\ell\rangle = \langle k, q, \ell, m_\ell|JM\rangle f(\beta, \tilde{\beta}, J, k, \ell). \tag{10.28}$$

This is profoundly useful for two reasons, both related to the fact that the function f does not depend on any of the projections q, m, or M. First, many matrix elements turn out to be zero because of the vanishing Clebsch-Gordan coefficient. Secondly, if one needs multiple matrix elements differing only by changes in the projections M, m_ℓ of q, one need only calculate for one set of projections, then obtain the remaining ones by taking ratios of Clebsch-Gordan coefficients.

To prove the Wigner-Eckart theorm, one first defines the state,

$$|\beta', k, \ell, j, m\rangle \equiv \sum_{q, m_\ell} \langle k, q, \ell, m_\ell|jm\rangle T^k_q|\beta, \ell, m_\ell\rangle. \tag{10.29}$$

This must rotate as an object with angular momentum J' and projection M'. Using the completeness of Clebsch-Gordan coefficients, one can also state that

$$T^k_q|\beta,\ell,m_\ell\rangle = \sum_{jm}\langle k,q,\ell,m_\ell|jm\rangle|\beta',k,\ell,j,m\rangle. \tag{10.30}$$

To prove the Wigner Eckart theorem, one must prove that the matrix element $\langle\tilde{\beta},J,M|\beta',k,\ell,j,m\rangle$ is (a) proportional to $\delta_{Mm'}$ (b) proportional to δ_{Jj} and (c) independent of M (aside from the delta function). Proving (a) and (b) rely on noting that the overlaps are independent of rotation if both the bra and ket rotate together,

$$\langle\tilde{\beta},J,M|\beta',k,\ell,j,m\rangle = \langle\tilde{\beta},J,M|R^{-1}(\vec{\alpha})R(\vec{\alpha})|\beta',k,\ell,j,m\rangle. \tag{10.31}$$

Under rotations about the z axis the bra and ket simply pick up a phase under rotation, e.g. $|\tilde{\beta},J,M\rangle \rightarrow |\tilde{\beta},J,M\rangle e^{iM\phi}$. Thus, if one averages the overlap over all azimuthal angles, the result will be proportional to δ_{Mm}. This proves (a). To prove (b), one can consider the average over all possible rotations of the three Euler angles. If one expresses the rotated states in terms of the $\mathcal{D}$ matrices, e.g. the bra and ket become

$$\begin{aligned} R(\vec{\omega})|\beta',j,m\rangle &= \sum_{m'}\mathcal{D}^j_{mm'}(\vec{\omega}))^*|\beta',j,m'\rangle, \\ \langle\tilde{\beta},J,M|R^{-1}(\vec{\omega}) &= \sum_{M'}\langle\tilde{\beta},J,M'|\mathcal{D}^J_{M'M}(\vec{\omega}). \end{aligned} \tag{10.32}$$

Taking the overlaps,

$$\begin{aligned} \langle\tilde{\beta},J,M|R^{-1}(\vec{\omega})R(\vec{\omega})|\beta',j,m\rangle &= \sum_{M'm'}\langle\tilde{\beta},J,M'|\mathcal{D}^J_{M'M}(\vec{\omega})\mathcal{D}^j_{mm'}(\vec{\omega}))^*|\beta',j,m'\rangle, \\ \langle\tilde{\beta},J,M|\beta',j,m\rangle &= \frac{1}{2J+1}\sum_{M'm'}\langle\tilde{\beta},J,M'|\beta',j,m'\rangle\delta_{J,j}\delta_{M'm'}\delta_{Mm} \\ &= \frac{1}{2J+1}\delta_{Jj}\delta_{Mm}\langle\tilde{\beta},J,M|\beta',Jm\rangle\sum_{m'M'}\delta_{m'M'} \\ &= \delta_{Jj}\delta_{Mm}\langle\tilde{\beta},J,M|\beta',JM\rangle. \end{aligned}$$

The second line invoked the orthogonality relations for the Wigner $\mathcal{D}$ functions, Eq. (10.20). Finally, to prove (c), one needs to see that, aside from the δ_{Mm} factor, there is no other M dependence, or equivalently that $\langle\tilde{\beta},J,M|\beta',k,\ell,j,M\rangle$ is independent of M. To see this consider the raising and lowering operators. One can see that

$$\begin{aligned} |\beta',k,\ell,J,M-1\rangle &= \frac{1}{\sqrt{J(J+1)-M(M-1)}}J_-|\beta',k,\ell,J,M\rangle, \\ |\tilde{\beta},J,M-1\rangle &= \frac{1}{\sqrt{J(J+1)-M(M-1)}}J_-|\tilde{\beta},J,M\rangle. \end{aligned} \tag{10.33}$$

Thus,

$$\begin{aligned}
&\langle\tilde{\boldsymbol{\beta}}, \boldsymbol{J}, \boldsymbol{M}-\mathbf{1}|\boldsymbol{\beta}', \boldsymbol{k}, \boldsymbol{\ell}, \boldsymbol{J}, \boldsymbol{M}-\mathbf{1}\rangle \\
&\quad= \frac{\mathbf{1}}{\boldsymbol{J}(\boldsymbol{J}+\mathbf{1})-\boldsymbol{M}(\boldsymbol{M}-\mathbf{1})}\langle\tilde{\boldsymbol{\beta}}, \boldsymbol{J}, \boldsymbol{M}|\boldsymbol{J}_+\boldsymbol{J}_-|\boldsymbol{\beta}', \boldsymbol{k}, \boldsymbol{\ell}, \boldsymbol{J}, \boldsymbol{M}\rangle \\
&\quad= \frac{\mathbf{1}}{\boldsymbol{J}(\boldsymbol{J}+\mathbf{1})-\boldsymbol{M}(\boldsymbol{M}-\mathbf{1})}\langle\tilde{\boldsymbol{\beta}}, \boldsymbol{J}, \boldsymbol{M}|\boldsymbol{J}_x^2+\boldsymbol{J}_y^2+\boldsymbol{J}_z|\boldsymbol{\beta}', \boldsymbol{k}, \boldsymbol{\ell}, \boldsymbol{J}, \boldsymbol{M}\rangle/\hbar^2 \\
&\quad= \frac{\mathbf{1}}{\boldsymbol{J}(\boldsymbol{J}+\mathbf{1})-\boldsymbol{M}(\boldsymbol{M}-\mathbf{1})}\langle\tilde{\boldsymbol{\beta}}, \boldsymbol{J}, \boldsymbol{M}|\boldsymbol{J}(\boldsymbol{J}+\mathbf{1})-\boldsymbol{M}(\boldsymbol{M}-\mathbf{1})|\boldsymbol{\beta}', \boldsymbol{k}, \boldsymbol{\ell}, \boldsymbol{J}, \boldsymbol{M}\rangle \\
&\quad= \langle\tilde{\boldsymbol{\beta}}, \boldsymbol{J}, \boldsymbol{M}|\boldsymbol{\beta}', \boldsymbol{k}, \boldsymbol{\ell}, \boldsymbol{J}, \boldsymbol{M}\rangle.
\end{aligned} \tag{10.34}$$

Putting (a), (b) and (c) together, Eq. (10.30) becomes

$$\begin{aligned}
\langle\tilde{\boldsymbol{\beta}}, \boldsymbol{J}, \boldsymbol{M}|\boldsymbol{T}_q^k|\boldsymbol{\beta}, \boldsymbol{\ell}, \boldsymbol{m}_\ell\rangle &= \sum_{jm}\langle\boldsymbol{k}, \boldsymbol{q}, \boldsymbol{\ell}, \boldsymbol{m}_\ell|jm\rangle\langle\tilde{\boldsymbol{\beta}}, \boldsymbol{J}, \boldsymbol{M}|\boldsymbol{\beta}', \boldsymbol{k}, \boldsymbol{\ell}, j, m\rangle \\
&= \langle\boldsymbol{k}, \boldsymbol{q}, \boldsymbol{\ell}, \boldsymbol{m}_\ell|\boldsymbol{J}\boldsymbol{M}\rangle\langle\tilde{\boldsymbol{\beta}}, \boldsymbol{J}, \boldsymbol{M}|\boldsymbol{\beta}', \boldsymbol{k}, \boldsymbol{\ell}, \boldsymbol{J}, \boldsymbol{M}\rangle,
\end{aligned} \tag{10.35}$$

and the final matrix element is independent of $\boldsymbol{M}$. To make it clear that the matrix element is independent of $\boldsymbol{M}$, the Wigner-Eckart theorem is expressed as

$$\langle\tilde{\boldsymbol{\beta}}, \boldsymbol{J}, \boldsymbol{M}|\boldsymbol{T}_q^k|\boldsymbol{\beta}, \boldsymbol{\ell}, \boldsymbol{m}_\ell\rangle = \langle\boldsymbol{k}, \boldsymbol{q}, \boldsymbol{\ell}, \boldsymbol{m}_\ell|\boldsymbol{J}\boldsymbol{M}\rangle\frac{\langle\tilde{\boldsymbol{\beta}}, \boldsymbol{J}, ||\boldsymbol{T}^{(k)}||\boldsymbol{\beta}, \boldsymbol{\ell}\rangle}{\sqrt{\mathbf{2}\boldsymbol{J}+\mathbf{1}}}. \tag{10.36}$$

The odd form with the two vertical bars carries no significance, aside from stating that the matrix element does not depend on the projection $\boldsymbol{M}$, but can depend on $\boldsymbol{J}$, $\boldsymbol{\beta}$, $\tilde{\boldsymbol{\beta}}$, $\boldsymbol{k}$ and $\boldsymbol{\ell}$. The choice of the $\sqrt{\mathbf{2}\boldsymbol{J}+\mathbf{1}}$ in the denominator is also arbitrary, but conventional. The matrix element with the double bars is known as the reduced matrix element. It is confusing as it implies that it is an object that one may calculate directly. This is not true. One must first calculate one of the matrix elements for a specific combination of $\boldsymbol{M}, \boldsymbol{m}_\ell$ and $\boldsymbol{q}$, then generate the reduced matrix element. Finally, from the reduced matrix element one can express all other matrix elements as the reduced matrix element multiplied by a Clebsch-Gordan coefficient.

The Wigner Eckart theorem is profound for two reasons.

1. An irreducible tensor operator $\boldsymbol{T}_q^k$ will not link two states of good angular momentum unless the Clebsch-Gordan coefficient coupling the angular momentum of the ket with the angular momentum of $\boldsymbol{T}$ to the angular momentum of the bra is not vanishing. Thus, a vector operator can not connect a $\boldsymbol{j}=\mathbf{2}$ state to a $\boldsymbol{j}=\mathbf{0}$ state. Also, the projections must add up zero, $\boldsymbol{q}+\boldsymbol{m}=\boldsymbol{m}'$.
2. If one needs to calculate $\langle\boldsymbol{\beta}', \boldsymbol{k}, \boldsymbol{j}, \boldsymbol{j}', \boldsymbol{m}'|\boldsymbol{T}_q^k|\boldsymbol{\beta}, \boldsymbol{j}, \boldsymbol{m}\rangle$ for some given set of $\boldsymbol{m}, \boldsymbol{m}'$ and $\boldsymbol{q}$, one can choose any values of $\boldsymbol{m}, \boldsymbol{m}'$ and $\boldsymbol{q}$ that might make calculation of the matrix element most simple then use the fact that the desired matrix element is the calculated matrix element multiplied by the ratio of the Clebsch-Gordan coefficients. For example if one needs to calculate a matrix element with $\boldsymbol{m}'=\mathbf{1},\ \boldsymbol{m}=-\mathbf{1}$ and $\boldsymbol{q}=\mathbf{2}$, one could simple calculate the matrix element with all three projections equal to zero, then multiply by the

ratio of Clebsch-Gordan coefficients.

$$\begin{aligned}\langle \beta', k, j, j', m' = 1 | T^k_{q=2} | \beta, j, m = -1 \rangle \\ = \langle \beta', k, j, j', m' = 0 | T^k_{q=0} | \beta, j, m = 0 \rangle \frac{\langle k, j, j', 1 | k, j, 2, -1 \rangle}{\langle k, j, j', 0 | k, j, 0, 0 \rangle}.\end{aligned} \tag{10.37}$$

This is handy because the integrals required to find the matrix element may be much easier to perform with the projections all set to zero, and if one needs to find the elements for many values of the projections, the matrix element must be calculated only once for each J, as ratios Clebsch-Gordan coefficients may be used to find all other matrix elements.

Example 10.4: Dipole Operators and the Wigner-Eckart Theorem
Suppose one needed to calculate the matrix elements, $\langle \tilde{\beta}, J' = 2, M' | x | \beta, J = 1, M \rangle$, for all M and M'.

1. For which values of M' are the elements zero?
2. If one calculated the matrix element for $\gamma \equiv \langle \tilde{\beta}, J' = 2, M' = 0 | z | \beta, J = 1, M = 0 \rangle$, express all the non-zero matrix elements above in terms of γ and Clebsch-Gordan coefficients.

Solution:
Using Eq. (10.14) one can write x, y and z in terms of irreducible tensor operators,

$$\begin{aligned} x &= \frac{1}{\sqrt{2}}(-T^1_1 + T^1_{-1}), \\ y &= \frac{i}{\sqrt{2}}(T^1_1 + T^1_{-1}), \\ z &= T^1_0. \end{aligned} \tag{10.38}$$

From this, one can see that from the fact that x either raises or lowers the projection,

$$\langle \tilde{\beta}, J' = 2, M' = 2 | x | \beta, J = 1, M = 0 \rangle = \langle \tilde{\beta}, J = 2, M = 2 | x | \beta, J = 1, M = -1 \rangle = 0, \tag{10.39}$$

$$\langle \tilde{\beta}, J' = 2, M' = 1 | x | \beta, J = 1, M = 1 \rangle = \langle \tilde{\beta}, M = 1 | x | \beta, J = 1, M = -1 \rangle = 0,$$

$$\langle \tilde{\beta}, J' = 2, M' = 0 | x | \beta, J = 1, M = 0 \rangle = 0,$$

$$\langle \tilde{\beta}, J' = 2, M' = -1 | x | \beta, J = 1, M = 1 \rangle = \langle \tilde{\beta}, M = -1 | x | \beta, J = 1, M = -1 \rangle = 0,$$

$$\langle \tilde{\beta}, J' = 2, M' = -2 | x | \beta, J = 1, M = 1 \rangle = \langle \tilde{\beta}, J = 2, M = -2 | x | \beta, J = 1, M = 0 \rangle = 0.$$

The non-zero elements are then given by one element you know, γ, and ratios of Clebsch-

Gordan coefficients.

$$\begin{aligned}
\langle\tilde{\beta}, J'=2, M'=2|x|\beta, J=1, M=1\rangle &= -\frac{\gamma}{\sqrt{2}}\frac{\langle k=1, q=1, J=1, M=1|J'=2, M'=2\rangle}{\langle k=1, q=0, J=1, M=0|J'=2, M'=0\rangle}, \qquad (10.40)\\
\langle\tilde{\beta}, J'=2, M'=1|x|\beta, J=1, M=0\rangle &= -\frac{\gamma}{\sqrt{2}}\frac{\langle 1,1,1,0|2,1\rangle}{\langle 1,0,1,0|2,0\rangle},\\
\langle\tilde{\beta}, J'=2, M'=0|x|\beta, J=1, M=-1\rangle &= -\frac{\gamma}{\sqrt{2}}\frac{\langle 1,1,1,-1|2,0\rangle}{\langle 1,0,1,0|2,0\rangle},\\
\langle\tilde{\beta}, J'=2, M'=0|x|\beta, J=1, M=1\rangle &= \frac{\gamma}{\sqrt{2}}\frac{\langle 1,-1,1,1|2,0\rangle}{\langle 1,0,1,0|2,0\rangle},\\
\langle\tilde{\beta}, J'=2, M'=-1|x|\beta, J=1, M=0\rangle &= \frac{\gamma}{\sqrt{2}}\frac{\langle 1,-1,1,0|2,-1\rangle}{\langle 1,0,1,0|2,0\rangle},\\
\langle\tilde{\beta}, J'=2, M'=-2|x|\beta, J=1, M=-1\rangle &= \frac{\gamma}{\sqrt{2}}\frac{\langle 1,-1,1,-1|2,-2\rangle}{\langle 1,0,1,0|2,0\rangle}.
\end{aligned}$$

The Wigner Eckart theorem explains many of the selection rules associated with atomic or nuclear transitions. In the dipole approximation the transition operator behaves as if it has $J = 1$. Thus, there are no transitions with $|\Delta J| > 1$ because the Clebsch-Gordan coefficients would then vanish.

Example 10.5: Wigner Eckart Practice
A group of particles is bound into a state $|\alpha, J, M\rangle$, where J and M reference the total angular momentum and α is the set of all other labels required to define the state.

1. Which of the matrix elements below might be non-zero?
   (a) $\langle\alpha', J'=2, M'=1|P_x^2 + P_y^2|\alpha, J=4, M=3\rangle$
   (b) $\langle\alpha', J'=2, M'=1|P_xP_y|\alpha, J=4, M=1\rangle$
   (c) $\langle\alpha', J'=2, M'=1|\sum_{ijk}\epsilon_{ijk}J_iR_jP_k|\alpha, J=2, M=1\rangle$
   (d) $\langle\alpha', J'=2, M'=1|P_x|\alpha, J=3, M=1\rangle$
2. With great effort you calculated the matrix element
   $$\mathcal{M} = \langle\alpha', J'=2, M'=0|P_x^2 + P_y^2 - 2P_z^2|\alpha, J=4, M=0\rangle.$$
   In terms of $\mathcal{M}$ and Clebsch-Gordan coefficients, express the following matrix elements
   (a) $\langle\alpha', J'=2, M'=1|P_x^2 + P_y^2 - 2P_z^2|\alpha, J=4, M=1\rangle$
   (b) $\langle\alpha', J'=2, M'=2|P_x^2 + P_y^2|\alpha, J=4, M=2\rangle$
   (c) $\langle\alpha', J'=2, M'=0|P_xP_y|\alpha, J=4, M=0\rangle$
   (d) $\langle\alpha', J'=2, M'=0|P_x^2 + P_y^2 - 2P_z^2|\alpha, J=2, M=0\rangle$

Solution:
1a) The operator can be written as a sum of T_0^2 and T_0^0 operators. It is zero becaue $M \neq M'$.
1b) The operator can be written in terms of T_1^2 and T_{-1}^2 operators. It is zero because the bra and ket have the same M.
1c) The operator is a scalar, and the bra and ket have the same J and M, so this can be non-zero.
1d) The operator can be written in terms of T_1^2 and T_{-1}^2 operators, so it must be zero because $M = M'$.
Next, we address the question 2a-2d. Using Eq.s (10.13) and (10.14) for guidance one can write

$$\begin{aligned} P_x^2 + P_y^2 - 2P_z^2 &= -2T_0^2, \\ P_x^2 + P_y^2 &= \frac{1}{3}\left[(P_x^2 + P_y^2 - 2P_z^2) + 2(P_x^2 + P_y^2 + P_z^2)\right] \\ &= -\frac{2}{3}T_0^2 + \frac{2}{3}T_0^0, \\ P_x P_y &= \frac{i}{\sqrt{6}}(T_{-1}^2 - T_1^2). \end{aligned}$$

Using the Wigner-Eckart theorem, first find one of the matrix elements of the irreducible tensor operator,

$$\langle \alpha', J' = 2, M' = 0 | T_{q=0}^{k=2} | \alpha, J = 4, M = 0 \rangle = -\frac{\mathcal{M}}{2}.$$

2a)

$$\frac{-\mathcal{M}}{2}\frac{\langle 2,0,4,1|2,1\rangle}{\langle 2,0,4,0|2,0\rangle}(-2)$$

2b) Note that $P_x^2 + P_y^2$ will be a combination of T_0^2 and a different irreducible operator T_0^0, but this last part gives zero because it a $k = 0$ operator can't couple with a $J = 4$ state to create a $J = 2$ state.

$$\frac{-\mathcal{M}}{2}\frac{\langle 2,0,4,2|2,2\rangle}{\langle 2,0,4,0|2,0\rangle}\left(\frac{-2}{3}\right)$$

2c) This is zero because $q + M \neq M'$
2d) The matrix element might be non-zero but it cannot be written in terms of $\mathcal{M}$ because $J = 2$ in the bra whereas $J = 4$ for $\mathcal{M}$.

10.8 Exercises

1. The $\Delta^{++,+,0,-}$ baryons have isospin 3/2 while the $\pi^{+,0,-}$ mesons form an isotriplet. Calculate the branching ratios of all four Δ decays into the corresponding $p\pi$ or $n\pi$ channels. (For instance, what fraction of the Δ^+s decay into $p\pi^0$ vs the $n\pi^+$ channels.)
2. The $S(975)$ meson is an isoscalar ($I = 0$), and decays into two pions. What fraction of the two-pion decays are expected to go into the neutral pion channel?

3. Write the Racah coefficient, $W(j_1, j_2, j_3, J; J_{12}, J_{23})$ which is defined by

$$\langle (j_1, j_2), J_{12}, j_3, J, M' | j_1, (j_2, j_3), J_{23}, J, M \rangle = \delta_{M,M'} \sqrt{(2J_{12}+1)(2J_{23}+1)} W(j_1, j_2, j_3, J; J_{12}, J_{23}),$$

in terms of Clebsch-Gordan coefficients.

4. For each operator, define a set (or sets) of irreducible tensor operators T^k_q, from which one can then define the given operator as a linear sum of the irreducible operators. (When defining a set, write down T^k_q for all possible q.)

   (a) z
   (b) p_x
   (c) x^2
   (d) $L_x L_y$

5. In terms of the Pauli matrices, find the rotation matrix $\mathcal{D}^{(j)}_{mm'}(\phi, \theta, \psi)$ for the case where $j = 1/2$, and ϕ, θ and ψ are Euler angles.

6. Using Eq.s (10.16) and (10.17) derive Eq. (10.18).

7. Assume one is calculating matrix elements for transitions from a d state to a s state via a quadrupole type coupling, and that one has performed an integral and found

$$I \equiv \langle \ell' = 0, m' = 0 | (z^2 - r^2/3) | \ell = 2, m = 0 \rangle.$$

   Given that one knows I, find

$$\langle \ell' = 0, m' = 0 | (x^2 - r^2/3) | \ell = 2, m \rangle$$

   for all five values of m, in terms of I. You can leave your answer in terms of Clebsch-Gordan coefficients.

8. The matrix element for the electromagnetic decay of an atomic d state with m_i to a p state with m_f is given by the matrix element,

$$\mathcal{M} \equiv \alpha \vec{\epsilon}^* \cdot \langle \ell = 1, m_f | \vec{r} | \ell = 2, m_i \rangle$$

   where $\alpha \vec{\epsilon}^* \cdot \vec{r}$ is the interaction responsible for the decay, and $\vec{\epsilon}$ is the polarization vector of the outgoing photon.

   Consider the intensity of RCP light that is emitted along the z axis. The polarization vector of such light can be written as $(1/\sqrt{2})(\hat{x} + i\hat{y})$. Find the RELATIVE intensities of such light for all 15 combinations of m_i and m_f. You can get Clebsch-Gordan coefficients from a table, e.g. https://en.wikipedia.org/wiki/Table_of_Clebsch-Gordan_coefficients, or use an on-line calculator, e.g. https://www.wolframalpha.com/input/?i=Clebsch-Gordan+calculator.

11 Fermions

When a system is comprised of a single particle, it does not matter whether it is a fermion or a boson. However, when multiple particles are present, behaviors can differ radically. The difference between fermions and bosons can be explained by a variety of perspectives.

1. When considering populations of single-particle levels, no more than one fermion of a given type and spin can be assigned to a particular level, while an arbitrary number of bosons can be placed in a given level. This constraint on fermions is known as the Pauli exclusion principle.
2. For identical fermions of the same spin in three single-particle states a, b, c, the multi-particle wave function $\psi_{abc}(x_1, x_2, x_3)$ must be completely anti-symmetric with respect any interchange of the three spatial coordinates or under interchange of any of the labels a, b, c. In contrast, the wave function for bosons must be symmetric under such interchanges.
3. Creation and destruction operators for fermions obey anti-commutation relations while those for bosons obey commutation relations.

Examples of fermions are electrons, quarks and neutrinos. Examples of bosons are photons and gluons. Composite particles made of an odd number of fermions, e.g. a proton which is made of three quarks, are also fermions. Composite particles made of an even number of fermions are bosons, e.g. a ^{12}C atom which is made of 6 protons, 6 neutrons and 6 electrons. All bosons have integral spin while all fermions have half-integral spin.

First, we will consider consequences of the Pauli exclusion principle. Later sections will presents discussions from the perspectives of anti-symmetrized wave functions or anti-commuting field operators.

11.1 The Spin Statistics Theorem

Consider the overlap of a state with N particles, where we refer to this N-particle state as $|\Phi\rangle$. One can take the overlap of this state with the state $\langle \vec{r}_1, s_1; \vec{r}_2, s_2; \cdots ; \vec{r}_N, s_N|\Phi\rangle$, where s_i is the spin of the i^{th} particle, which is at $\vec{r}_i$. If one writes the same overlap, but with a set of two indices permuted, e.g. $\vec{r}_4, s_4 \leftrightarrow \vec{r}_7, \vec{s}_7$, the new overlap will be the same but with a relative sign of ± 1. The permutation operator $\mathcal{P}$, which in this example permutes the $i = 4$ and $i = 7$ particles, cannot change the physics if the particles are identical. Thus, it must commute with the Hamiltonian, and because $\mathcal{P}^2$ must return to the original state, its eigenvalues must be ± 1. The spin-statistics theorem states that for half-integral-spin ($s = 1/2, 3/2, 5/2 \cdots$) particles the eigenvalue is -1, and for integral spins the eigenvalue is unity. Thus, wave functions for integral-spin particles (bosons) are symmetric and the wave functions for half-integral particles (fermions) are anti-symmetric under permutation.

To motivate the spin-statistic theorem, one can consider two spin half particles, each with spin $\langle S_z \rangle = 1/2$. The two particles are in the $x - y$ plane at positions at opposite sides of a circle centered at the origin. If one were to rotate about the $z-$axis by a phase ϕ =180 degrees, the

spin wave function for each particle would change by a phase,

$$e^{i\sigma_z\phi/2} = \cos(\phi/2) + i\sigma_z \sin(\phi/2). \tag{11.1}$$

For spin-up particles, this introduces a phase factor of i when the angle of rotation is 180 degrees. A second phase factor of i comes from rotating the second particle, resulting in an overall phase factor of -1. The two particles are now in the same state but permuted. Thus, the permutation is equivalent to this 180 degree rotation and has a phase factor of -1 relative to the unrotated, or un-permuted, state. For spin-down fermions one would have an overall phase factor of -1. For bosons, rotating by 180 degrees yields a factor of ± 1 for each particle, which gives an overall factor of $+1$. Thus, half-integral-spin particles should be anti-symmetric under permutation, while permutation of integral-spin particles should yield the same overall state, without a sign change. For the two-particle wave functions of fermions,

$$\begin{aligned} \psi_{ab}(\vec{r}_1, \vec{r}_2) &= \langle \vec{r}_1, \vec{r}_2 | a, b\rangle \\ &= -\psi_{ab}(\vec{r}_2, \vec{r}_1) = -\langle \vec{r}_2, \vec{r}_1 | a, b\rangle \\ &= -\psi_{ba}(\vec{r}_1, \vec{r}_2) = -\langle \vec{r}_1, \vec{r}_2 | b, a\rangle \\ &= \psi_{ba}(\vec{r}_2, \vec{r}_1) = \langle \vec{r}_2, \vec{r}_1 | b, a\rangle. \end{aligned} \tag{11.2}$$

The anti-symmetry property applies to either the $\vec{r}_i$ labels or to the states a, b.

If the two-particle wave function is anti-symmetric, the two particles cannot occupy the same single-particle state. If one uses field operators to create or destroy particles, the operators must obey anti-commutation relations, rather than the commutation relations used in Sec. 9.2, if the particles are fermions. Fermi field operators are discussed in Sec. 11.5.

11.2 The Pauli-Exclusion Principle

The wave-function $\psi_{a,s_a;b,s_b}(\vec{r}_1, s_1; \vec{r}_2, s_2)$ must be anti-symmetric under interchange of a and b. So if $a = b$ the wave function must be zero. If one considers the set of eigenstates of an external potential $V(\vec{r})$, then each single-particle eigenstate can only host a single particle if the spins $a = b$ and $s_a = s_b$. If the spins, s_a and s_b are different, or if $a \neq b$, then one can write an anti-symmetrized version of the wave function. For that reason for each single-particle eigenstate of an external potential, only one fermion can occupy a given eigenstate, or orbital. If there is spin-degeneracy, then the same energy level might accommodate several different fermions if they have different angular momentum or spin. This constraint is prominent in understanding electronic structures of multi-electron atoms of in understanding how neutrons and protons can arrange themselves to form nuclei.

Example 11.1: Filling the Harmonic Oscillator with Fermions
Here, we consider the filling of harmonic oscillator energy levels with fermions. This simple picture provides a surprising amount of insight into the structure of light nuclei.
First, we must understand the degeneracy of single-particle levels in a 3-d harmonic oscillator. Each level has an energy of $(N + 3/2)\hbar\omega$. Considering the problem in a Cartesian basis, there are $N_\perp + 1$ ways to arrange $n_x + n_y = N_\perp$ (Here, we consider only spin-up fermions). The

number of ways to arrange n_x, n_y and n_z to add up to N is

$$d(N) = \sum_{N_\perp=0}^{N} N_\perp + 1 = \frac{(N+1)(N+2)}{2}.$$

Thus, there is one combination of n_x, n_y and n_z to get $N = 0$, three combinations to get $N = 1$, six to get $N = 2$, etc.
Next, we wish to calculate the number of states of a specific N that have a given orbital angular momentum ℓ. This will come in handy when we consider the spin-orbit coupling on top of the harmonic oscillator structure. The solution will be that the states with excitation N are accounted for with one $\ell = N$ multiplet, one $\ell = N-2$ multiplet, one $\ell = N-4$ multiplet, etc. To prove that this is the case, we make an inductive proof assuming it is true for $N-2$. For every state with excitation $N-2$ there is a state with the same angular momentum with excitation N which is reached by applying the additional operator $(a_x^\dagger a_x^\dagger + a_y^\dagger a_y^\dagger + a_z^\dagger a_z^\dagger)$ which is a rotational scalar. There are $(N-1)N/2$ such states. Furthermore, there must be at least one $\ell = N$ multiplet because the state $(a_x^\dagger + i a_y^\dagger)^N |0\rangle$ transforms like part of a $\ell = N$ multiplet. But one $\ell = N$ multiplet with degeneracy $(2N+1)$, and the $(N-1)N/2$ states accounted for by counting the states with excitation $N-2$ completely account for all the states with excitation N,

$$\frac{(N+1)(N+2)}{2} = \frac{(N-1)N}{2} + (2N+1).$$

Thus, increasing the excitation by $2\hbar\omega$ adds one more multiplet with $\ell = N$. As an example, for $N = 4$, there are multiplets with $\ell = 4, 2, 0$ with degeneracy $9 + 5 + 1 = 15 = 5 \cdot 6/2$.
If no spin-orbit terms were present, adding neutrons to a harmonic oscillator would lead to shell closures with neutron numbers equal to 2, 8, 20, 40 and 70. We now consider filling all the $N \leq 3$ shells, which requires 40 neutrons. Note that these numbers accounted for both spin-up and spin-down levels. To calculate this number for $N \leq 3$ one would consider that there is $1 \cdot 2/2 = 1$ way to arrange n_x, n_y, n_z for $N = 0$ or $E = 3\hbar\omega/2$, $2 \cdot 3/2 = 3$ ways to make arrangements for $E = 5\hbar\omega/2$ single-particle levels, $3 \cdot 4/2 = 6$ single-particle levels with $E = 7\hbar\omega/2$ and $4 \cdot 5/2 = 10$ single-particle levels with $E = 9\hbar\omega/2$. This sums to $1 + 3 + 6 + 10 = 20$ single-particle levels with $E \leq 9\hbar\omega/2$. Because each single-particle level can have either spin-up or spin-down, this gives 40 single-particle levels for $N \leq 3$, or equivalently $E \leq 9\hbar\omega/2$, for spin $1/2$ particles. Thus, one could place up to 40 neutrons into the 40 single-particle levels without having any particle's energy exceed $9\hbar\omega/2$.
Adding the spin-orbit term adjusts the single-particle energies by an amount

$$E_{\text{s.o.}} = -\beta \vec{\ell} \cdot \vec{s} = -\frac{\beta\hbar^2}{2}\left[j(j+1) - \ell(\ell+1) - s(s+1)\right].$$

For every value of ℓ there are two values of j, $j = \ell \pm 1/2$. Thus, one finds the energy levels by first labeling the states by N and ℓ. Then splitting each level into it's two values of j and finding its energy,

$$E = (N + 3/2)\hbar\omega - \frac{\beta\hbar^2}{2}\left[j(j+1) - \ell(\ell+1) - s(s+1)\right] - \alpha\ell(\ell+1).$$

An additional term, proportional to L^2, was also added to phenomenologically better reproduce levels. This is known as the Nilsson model, https://en.wikipedia.org/wiki/Nilsson_model. The shell structure changes due to the spin-orbit interaction, and the numbers of neutrons required to reach a condition where there is a large gap are called magic numbers. In nuclei, the magic numbers are 2, 8, 20, 28, 50, 82 and 126. Nuclei are considered doubly magic if both the neutron and proton numbers are magic numbers. Examples are ^{4}He, ^{16}0, $^{40}_{20}$Ca, ^{48}Ca, $^{56}_{28}$Ni, ^{48}Ni, ^{78}Ni, $^{100}_{50}$Sn, ^{132}Sn and $^{208}_{82}$Pb. The magic numbers of nuclei were only explained by assuming an anomalously large spin-orbit coupling. The reason for this surprisingly large coupling will be explained later in the semester when the Dirac equation is presented.

11.3 Fermi Gases and Neutron Stars

Consider a large number N of fermions of spin s and mass m in a large box of volume V. Here, the box is defined by a confining potential that is zero inside, and infinite outside, the box. If the fermions are placed in the lowest levels consistent with the Pauli exclusion principle, the highest single-particle energy is known as the Fermi energy ϵ_f and the momentum of that state is known as the Fermi momentum p_f. The density is a function of the Fermi momentum,

$$\begin{aligned} N &= (2s+1)\frac{V}{(2\pi\hbar)^3}\int^{p<p_f} d^3p, \\ n &= \frac{N}{V} = \frac{(2s+1)}{6\pi^2\hbar^3}p_f^3, \end{aligned} \tag{11.3}$$

where s is the intrinsic spin of the fermion, e.g. $1/2, 3/2 \cdots$. Note that the Fermi momentum is determined by the density and does not depend on the particle's mass, whereas the Fermi energy does depend on the mass, $\epsilon_f = p_f^2/(2m)$. For this course, we use the term "Fermi energy" to refer to the kinetic energy of a particle at the top of the Fermi surface. The term Fermi surface refers to the surface in momentum space of the sphere defining the occupied region of momentum space. Contributions to the energy from the mass or potential also play a role, but the term "Fermi energy" will only refer to the kinetic part of the energy. When reading the literature, one needs to be careful in understanding how the author is using the term.

One example where the Fermi energy plays a pivotal role is in neutron stars. Due to beta decays neutrons can change into protons through the emission of an electron and a neutrino,

$$n \rightarrow p + e + \bar{\nu}, \;\; p + e \rightarrow n + \nu. \tag{11.4}$$

Because neutrinos can exit the star due to their negligible masses and small cross-sections, they need not be considered with respect to conservation laws. However, baryon number (the net number of neutrons and protons) and electric charge must be conserved, and in fact, the net electric charge density must be zero. Thus, beta decays, and inverse beta decays, can proceed as long as the following constraints are met,

$$\begin{aligned} n_n + n_p &= n_B, \\ n_e &= n_p, \end{aligned} \tag{11.5}$$

where n_n, n_p and n_e are the neutron, proton and electron densities respectively and n_B is the baryon density. Beta decays will proceed until the energy is minimized for the given value of the baryon density n_B. Neutrons are more massive than protons by an amount $(m_n - m_p)c^2 = \mathbf{1.30}$ MeV which more than accounts for the electron's mass of 0.511 MeV$/c^2$. When the energy is minimized, changing a neutron to a proton plus electron must leave the energy unchanged if the particles are added and removed from the top of the Fermi surfaces. If net energy changes when undergoing such a reaction, reactions would proceed whichever way lowers the energy. These reactions would emit either a neutrino or anti-neutrino of vanishing energy to bring the energies into balance. For example if $\epsilon_f^{(p)} + \epsilon_f^{(e)} + m_p + m_e > \epsilon_f^{(n)} + m_n$, one could lower the energy by the reaction $p + e \rightarrow n + \nu$. This reaction would proceed until enough protons and electrons were taken from their Fermi surfaces, and enough neutrons were added to the top of the neutron Fermi sea until the following condition was met,

$$\epsilon_f^n + \mathbf{1.30\ MeV} = \epsilon_f^e + \epsilon_f^p + \mathbf{0.511\ MeV}. \tag{11.6}$$

Combining this equality with the two constraints in Eq. (11.5) allows one to find all three Fermi momenta in terms of n_B.

At very low density, the fact that neutrons are heavier than protons by 1.30 MeV$/c^2$ leads to protons being more populous than neutrons even though each proton must be accompanied by an electron with a mass of $\mathbf{0.511}$ MeV$/c^2$. Therefore, in stars like our sun protons are much more numerous than neutrons. At high density the electron's Fermi energy plays the pivotal role. Because electrons and protons have the same density, they have the same Fermi momentum, but the electron's Fermi energy is much higher due to its lighter mass. For a proton,

$$\epsilon_f^{(\mathrm{p})} = \sqrt{m_p^2 + p_f^2} - m_p \approx p_f^2/(2m_p), \tag{11.7}$$

which would be smaller if the protons mass were even larger. Electrons are lighter and thus have a much higher Fermi energy than protons, and in fact the densities tend to be so high that the electron's Fermi momenta well exceed the mass. The electron energies are then highly relativistic,

$$\epsilon_f^{(\mathrm{e})} = \sqrt{m_e^2 + p_f^2} - m_e \approx p_f. \tag{11.8}$$

Thus, at high density the system must try to reduce the electron density which in turn requires a reduction in the proton density in order to maintain charge neutrality. This results in a large excess of neutrons in order to achieve the required baryon density. This is certainly the case for the interior of neutron stars where densities are of the order 0.1 baryons per cubic fm. The resulting electron Fermi energy is on the order of 10 MeV as is the resulting neutron Fermi energy. At these densities the electrons are highly relativistic, because the electron Fermi energy is much higher than the electron mass, while the protons and neutrons, whose masses are approximately 939 MeV$/c^2$, remain largely non-relativistic.

Example 11.2: Zero-Point Surface Energy for Fermions
It costs energy to divide a piece of metal (which is approximated here as a non-interacting electron gas) into two pieces. The associated surface energy, energy per surface area, has a component deriving from the penalty associated with the kinetic energies of the particles

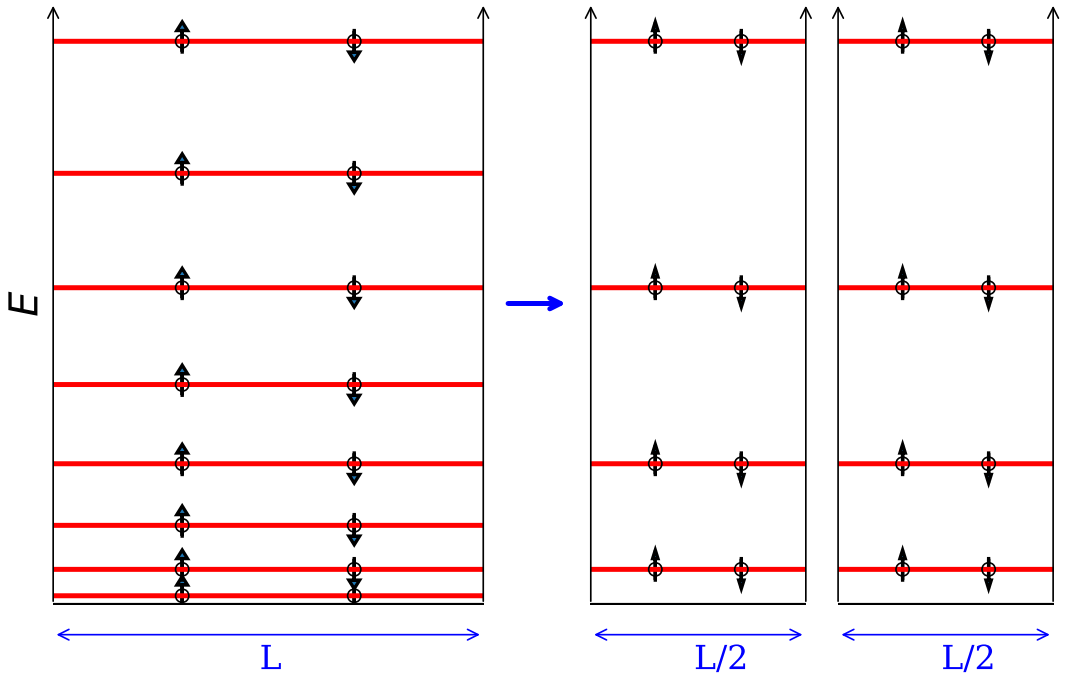


Figure 11.1: For Example 11.2: Energy levels of a one-dimensional infinite square well of length L are displayed in the left figure, while the levels of two wells of length $L/2$ are displayed in the two right panels. Each level can hold one fermion of a given spin. By splitting the well in two, the energy penalty is equivalent to moving fermions from the $1, 3, 5, 7 \cdots$ levels to the $2, 4, 6, 8 \cdots$ levels. Because half the fermions were moved up by one step in the ladder, the net energy penalty is $\epsilon_f/2$ for each spin in the limit of a large system where there are many levels.

known as the zero-point energy. To understand the source of this energy, we first consider the one-dimensional case. Consider a box of length L, which is divided into two boxes, each of length $L/2$. The initial energy levels of the larger box and the energy levels of the two smaller boxes are illustrated in Fig. 11.1. The levels are described by,

$$E_n = \frac{\hbar^2 k_n^2}{2m}, \quad k_n L = n\pi, \quad n = 1, 2, 3, \cdots$$

But after splitting the box the energy levels are given by

$$E_n = \frac{\hbar^2 k_n^2}{2m}, \quad k_n L = n\pi, \quad n = 2, 4, 6, \cdots,$$

with the new levels being able to hold twice as many particles because there are now two boxes.
Thus, in one dimension the energy penalty can be thought of as arising from forcing half the particles to move up by one rung on the energy diagram ladder. Neglecting spin, if all the particles had moved up one rung, the net penalty would be the Fermi energy, but because half the particles moved up, the penalty is

$$\Delta E = \epsilon_f/2.$$

The energy penalty would then be doubled to account for the two spins.
In three dimensions, there are three quantum numbers, n_x, n_y and n_z, corresponding to k_x, k_y and k_z.

$$k_x L_x = n_x \pi, \;\; k_y L_y = n_y \pi, \;\; k_z L_z = n_z \pi.$$

If the box is divided in two along the x axis, $L_x \rightarrow L_x/2$, the penalty for all levels with a fixed n_y and n_z is,

$$\delta E(n_y, n_z) = \frac{1}{2} \frac{\hbar^2 (k_f^2 - k_y^2 - k_z^2)}{2m}.$$

This result follows from the fact that half the particles moved up a notch on the n_x ladder, with the top energy being such that the overall energy (including the y and z components) adds up to the Fermi energy.
One can now find the entire energy cost by summing over n_y and n_z.

$$\begin{aligned}
\Delta E &= \frac{L_y L_z}{(2\pi)^2} \int_{|k|<k_f} dk_y dk_z \delta E(n_y, n_z) \\
&= \frac{\hbar^2 L_y L_z}{8m\pi} \int_0^{k_f} k dk \, (k_f^2 - k^2) \\
&= L_y L_z \frac{\hbar^2 k_f^4}{32 m \pi} \\
&= L_y L_z \frac{k_f^2}{16\pi} \epsilon_f.
\end{aligned}$$

In the second step, the variables k_x and k_y were transformed to polar coordinates, $k = \sqrt{k_x^2 + k_y^2}$. Dividing by $2L_y L_z$ (as extra area is added to both sides of the split) gives the surface energy. One would also multiply the result by the spin degeneracy.

$$\frac{\Delta E}{A} = (2s+1) \frac{k_f^2}{32\pi} \epsilon_f.$$

In real systems the surface energy is of this order. It is lessened by the fact that the electron density is not sharply cut off at the surface, but dies out smoothly. An additional positive contribution can result from finite-range attractive interactions between the particles. In nuclear physics, the above expression over-predicts the surface energy of nuclear matter by approximately 25%.

Example 11.3: Magnetization of a Fermi Gas
A magnetic field lowers the energies of those particles whose spin aligns with the magnetic field, while raising the energies of those that are anti-parallel. For a magnetic field pointing in the z direction, this results in a surplus of spin-up particles vs spin-down particles. The magnetic energy is

$$H_B = -g\mu_B \frac{\vec{S}}{\hbar} \cdot \vec{B},$$

where $\mu_B = e\hbar/2mc$ is the Bohr magneton, and for electrons $g = 2$. For a gas of non-interacting electrons of number density n_e, find the net spin, $n_\uparrow - n_\downarrow$ for a small magnetic field. Also, find the magnetic susceptibility, $\chi \equiv \mu_B(n_\uparrow - n_\downarrow)/B$.
Solution: The difference between the spin-up and spin-down energies is

$$\Delta E = \frac{e\hbar B}{mc}.$$

In terms of the two Fermi momenta, the densities are

$$\begin{aligned} n_\uparrow &= \frac{1}{6\pi^2}k_\uparrow^3 \ , \\ n_\downarrow &= \frac{1}{6\pi^2}k_\downarrow^3, \\ n_e &= n_\uparrow + n_\downarrow. \end{aligned}$$

The Fermi energies differ by ΔE,

$$\frac{\hbar^2 k_\downarrow^2}{2m} - \frac{\hbar^2 k_\uparrow^2}{2m} = \Delta E.$$

For small fields, we can write the expressions above in terms of $\Delta k = k_\uparrow - k_\downarrow$, and if the original Fermi momentum was k_f, then $k_\uparrow = k_f + \Delta k/2$, $k_\downarrow = k_f - \Delta k/2$.

$$\begin{aligned} \Delta E &= \frac{\hbar^2 k_f}{m}\Delta k, \\ n_\uparrow - n_\downarrow &= \frac{3k_f^2}{6\pi^2}\Delta k, \\ &= \frac{k_f^2}{2\pi^2}\frac{m}{\hbar^2 k_f}\Delta E \\ &= \frac{mk_f}{2\pi^2\hbar^2}\Delta E \\ &= \frac{ek_f}{2\pi^2 c\hbar}B \\ &= \frac{e}{2\pi^2 c\hbar}(3\pi^2 n_e)^{1/3}B. \end{aligned}$$

The magnetic moment of the electron is μ_B, so the magnetic moment density is

$$\begin{aligned} M &= \mu_B(n_\uparrow - n_\downarrow) \\ &= \frac{e\mu_B}{2\pi^2 c\hbar}(3\pi^2 n_e)^{1/3}B. \end{aligned}$$

The magnetic susceptibility is the ratio of the induced magnetic moment density to a small applied field,

$$\chi = \frac{e\mu_B}{2\pi^2 c\hbar}(3\pi^2 n_e)^{1/3}.$$

11.4 Multi-Particle Symmetrization and Slater Determinants

Consider a multi-particle state $|\alpha_1, \alpha_2 \cdots\rangle$ of distinguishable particles, where the ordering $\alpha_1, \alpha_2, \cdots$ signifies that the first particle is in the state α_1, the second in α_2, etc. The permutation operator $\mathcal{P}_{ij}$ is defined by

$$\mathcal{P}(12)|\alpha_1, \alpha_2, \alpha_3, \cdots\rangle = |\alpha_2, \alpha_1, \alpha_3, \cdots\rangle. \tag{11.9}$$

If any two particles are identical, and indistinguishable, the multi-particle state $|\phi\rangle$ should be an eigenstate of the permutation operator,

$$\mathcal{P}(12)|\alpha_1, \alpha_2, \alpha_3, \cdots\rangle = |\alpha_2, \alpha_1, \alpha_3, \cdots\rangle = p|\alpha_1, \alpha_2, \alpha_3, \cdots\rangle. \tag{11.10}$$

If the permuted state differed from the un-permuted state, aside from a phase factor, the observables related to that state might change. This should not be possible if the particles are identical. All states of the system should be eigenstates of the permutation operator $\mathcal{P}(i,j)$,

$$\mathcal{P}(i,j)|\phi\rangle = (\pm 1)|\phi\rangle. \tag{11.11}$$

Only ± 1 are possible eigenvalues of the permutation operator because applying the same permutation operator twice yields unity, $\mathcal{P}^2(i,j) = 1$. In order for the state to be an eigenstate of all permutations, an N-particle state of all identical particles must be either totally symmetric (bosons) or totally anti-symmetric (fermions).

$$|\phi\rangle = \frac{1}{\sqrt{N!}} \sum_{\text{perm.s } p} (\pm 1)^{n_p} \mathcal{P}_p |\alpha, \beta, \gamma \cdots\rangle, \tag{11.12}$$

where n_p is the number of pairwise permutations required to make the desired permutation. For fermions, no two states, e.g. α and β can be identical. As an example, the three particle antisymmetric state is

$$|\phi\rangle = \frac{1}{\sqrt{6}} \left\{ |\alpha, \beta, \gamma\rangle - |\beta, \alpha, \gamma\rangle - |\gamma, \beta, \alpha\rangle - |\alpha, \gamma, \beta\rangle + |\beta, \gamma, \alpha\rangle + |\gamma, \alpha, \beta\rangle \right\}. \tag{11.13}$$

Note that the fifth and sixth orderings required two pairwise permutations, whereas the second, third and fourth orderings could be generated from a single pair-wise switch of labels. Similarly, one can do the same for a state with coordinate-space labels,

$$|\phi'\rangle = \frac{1}{\sqrt{6}} \left\{ |\vec{r}_i, \vec{r}_j, \vec{r}_k\rangle - |\vec{r}_j, \vec{r}_i, \vec{r}_k\rangle - |\vec{r}_k, \vec{r}_j, \vec{r}_i\rangle - |\vec{r}_i, \vec{r}_k, \vec{r}_j\rangle + |\vec{r}_j, \vec{r}_k, \vec{r}_i\rangle + |\vec{r}_k, \vec{r}_i, \vec{r}_j\rangle \right\}. \tag{11.14}$$

Taking the overlap of $\langle\phi'|\phi\rangle$ gives the 3-particle wave function. The 36 terms reduce to six terms because so many are redundant,

$$\begin{aligned} \langle\phi'|\phi\rangle = \phi_{\alpha,\beta,\gamma}(\vec{r}_i, \vec{r}_j, \vec{r}_k) \\ = &\phi_\alpha(\vec{r}_i)\phi_\beta(\vec{r}_j)\phi_\gamma(\vec{r}_k) - \phi_\beta(\vec{r}_i)\phi_\alpha(\vec{r}_j)\phi_\gamma(\vec{r}_k) \\ &- \phi_\gamma(\vec{r}_i)\phi_\beta(\vec{r}_j)\phi_\alpha(\vec{r}_k) - \phi_\alpha(\vec{r}_i)\phi_\gamma(\vec{r}_j)\phi_\beta(\vec{r}_k) \\ &+ \phi_\beta(\vec{r}_i)\phi_\gamma(\vec{r}_j)\phi_\alpha(\vec{r}_k) + \phi_\gamma(\vec{r}_i)\phi_\alpha(\vec{r}_j)\phi_\beta(\vec{r}_k). \end{aligned} \tag{11.15}$$

This wave function is anti-symmetric with respect to the pairwise interchange of any of the three indices α, β, γ, and is also anti-symmetric with respect to the pairwise interchange of any of three spatial indices $\vec{r}_i, \vec{r}_j, \vec{r}_k$. Ignoring spin, the overlap of such a state in coordinate space, i.e. for any number of particles, n, the multi-particle wave function, $\phi_{\alpha_1 \cdots \alpha_n}(\vec{r}_1 \cdots \vec{r}_n)$ must also be an eigenstate of the permutation operator, and this must be true whether one is permuting the α_i indices or the $\vec{r}$ indices.

Notation in terms of bras and kets can be confusing. Above, the state $|\alpha, \beta, \gamma\rangle$ denoted a state of three distinguishable particles, where the first particle occupied the level α, the second occupied β and the third γ. Given that the particles are actually indistinguishable, from here on out, as is nearly always the case in the literature, the state $|\alpha, \beta, \gamma\rangle$ will simply refer to the fact that α, β and γ are occupied and will not imply a specific ordering. After all, it makes no sense to state"first particle in in level α, second is in β, $\cdots$", because there is no way to distinguish which particle is first, second or third. Using fermionic field operators to define states allows one to to avoid this confusion, and is the subject of the next section.

If the single particle wave functions $\phi_\alpha, \phi_\beta, \cdots$ are orthonormal, the multi-particle wave function will also be normalized,

$$\frac{1}{N!} \int d^3r_1 \cdots d^3r_n |\phi_{\alpha_1 \cdots \alpha_n}(\vec{r}_1 \cdots \vec{r}_n)|^2 = 1. \tag{11.16}$$

The $1/N!$ factor accounts for the fact that integrating over all positions multiply counts for any given configuration of $\vec{r}_i$. One can express the multi-particle wave function in terms of normalized single-particle wave functions, $\phi_\alpha(\vec{r})$, in the form of a determinant

$$\phi_{\alpha_1 \cdots \alpha_n}(\vec{r}_1 \cdots \vec{r}_n) = \begin{vmatrix} \phi_{\alpha_1}(\vec{r}_1) & \phi_{\alpha_1}(\vec{r}_2) & \cdots & \phi_{\alpha_1}(\vec{r}_n) \\ \phi_{\alpha_2}(\vec{r}_1) & \phi_{\alpha_2}(\vec{r}_2) & \cdots & \phi_{\alpha_2}(\vec{r}_n) \\ \cdots & \cdots & \cdots & \cdots \\ \phi_{\alpha_n}(\vec{r}_1) & \phi_{\alpha_n}(\vec{r}_2) & \cdots & \phi_{\alpha_n}(\vec{r}_n) \end{vmatrix}. \tag{11.17}$$

The expression of the anti-symmetric N-particle wave function as a Slater Determinant in Eq. (11.17) is often expressed with a pre-factor of $1/\sqrt{N!}$. When this convention is chosen, the completeness/normalization condition in Eq. (11.16) is simultaneously altered by removing the prefactor $1/N!$.

This becomes more complicated when one includes spin, or isospin indices. In fact, one might have a wave function that is symmetric in coordinate space, but anti-symmetric in spin, or vice versa. Such cases will be discussed in the next chapter.

The Slater determinant enforces the Pauli exclusion principle. If any of the indices, $\alpha_1 \cdots \alpha_n$, are identical, the multi-particle wave function vanishes. The determinant also vanishes in the limit that any two positions become equal. It should be emphasized that not all multi-particle wave functions can be written in terms of a Slater determinant. If one were considering distinguishable particles, the multi-particle wave function would be generally of the form $\phi_{\alpha_1 \cdots \alpha_n}(\vec{r}_1 \cdots \vec{r}_n)$. Assuming that ϕ factorizes into a product of single-partcle wave functions represents a strong assumption. Similarly, although the Slater determinant is not a simple product wave function, the assumption that it can be written as a sum of simply permuted product wave functions represents an equally strong assumption. In fact, this assumption is the foundation of the Hartree-Fock approximation, which is discussed in the next chapter.

Example 11.4: Writing Three-Particle Wave Function with Slater Determinants
Imagine one has 3 electrons in a Coulomb potential. Two of the electrons are in the $1s$ state. The radial wave function for this state is noted by $\phi_{1s}(r)$. The third electron is in the $2p$ level and has spin-up, where the radial wave function is labeled $\phi_{2p}(r)$. Write the three-particle wave function, where particles are at positions $\vec{r}_1, \vec{r}_2$ and $\vec{r}_3$ with spin projections $s_1, s_2 = \downarrow$ and s_3 respectively.
Solution: Use the Slater Determinant, with the labels being $(s,\uparrow)$, $(s,\downarrow)$ and $(p,\uparrow)$.

$$\begin{aligned}
&\Psi((\vec{r}_1,s_1),(\vec{r}_2,s_2),(\vec{r}_3,s_3)) \\
&\quad = \begin{vmatrix} \phi_{1s,\uparrow}(\vec{r}_1,s_1) & \phi_{1s,\uparrow}(\vec{r}_2,s_2) & \phi_{1s,\uparrow}(\vec{r}_3,s_3) \\ \phi_{1s,\downarrow}(\vec{r}_1,s_1) & \phi_{1s,\downarrow}(\vec{r}_2,s_2) & \phi_{1s,\downarrow}(\vec{r}_3,s_3) \\ \phi_{2p,\uparrow}(\vec{r}_1,s_1) & \phi_{2p,\uparrow}(\vec{r}_2,s_2) & \phi_{2p,\uparrow}(\vec{r}_3,s_3) \end{vmatrix} \\
&= \phi_{1s,\uparrow}(\vec{r}_1,s_1)\phi_{1s,\downarrow}(\vec{r}_2,s_2)\phi_{2p,\uparrow}(\vec{r}_3,s_3) - \phi_{1s,\uparrow}(\vec{r}_1,s_1)\phi_{1s,\downarrow}(\vec{r}_3,s_3)\phi_{2p,\uparrow}(\vec{r}_2,s_2) \\
&\quad - \phi_{1s,\uparrow}(\vec{r}_2,s_2)\phi_{1s,\downarrow}(\vec{r}_1,s_1)\phi_{2p,\uparrow}(\vec{r}_3,s_3) + \phi_{1s,\uparrow}(\vec{r}_2,s_2)\phi_{1s,\downarrow}(\vec{r}_3,s_3)\phi_{2p,\uparrow}(\vec{r}_1,s_1) \\
&\quad + \phi_{1s,\uparrow}(\vec{r}_3,s_3)\phi_{1s,\downarrow}(\vec{r}_1,s_1)\phi_{2p,\uparrow}(\vec{r}_2,s_2) - \phi_{1s,\uparrow}(\vec{r}_3,s_3)\phi_{1s,\downarrow}(\vec{r}_2,s_2)\phi_{2p,\uparrow}(\vec{r}_1,s_1).
\end{aligned}$$

The three states, $(1s,\uparrow)$, $(1s,\downarrow)$ and $(2p,\uparrow)$ were chosen to be different. If any two of the states were the same, e.g. replace $(1s,\downarrow)$ with $(1s,\uparrow)$ one can see that the resulting three-particle wave function would vanish. Similarly, if one chooses to evaluate the wave functions with two of the particles at the same position and the same spin, e.g. $(\vec{r}_1,s_1) = (\vec{r}_2,s_2)$, the three-particle wave function would also vanish. But, it does not vanish if two positions are the same, but with different spins, e.g. $\vec{r}_1 = \vec{r}_2, s_1 \neq s_2$.

11.5 Fermi Creation and Destruction Operators

Writing states as a sum over various permutations is a rather clumsy way to consider the Fermi nature of the particles as it requires assuming there is a "first" particle, "second" particle, and so on, then include $N!$ terms in the various expressions. The algebra of anti-commuting creation and destruction operators offers a more natural means to incorporate anti-symmetrization. With this formalism, a state can be noted simply by their labels with no mention of permutations. Matrix elements are then calculated according to the algebra of the creation and destruction operators, which account for the symmetrization or anti-symmetrization. This approach allows one to address systems with many fermions, even in the macroscopic limit.

Fermi creation and destruction operators obey the algebra

$$\begin{aligned}
\{a_\alpha, a_\beta^\dagger\} &= \delta_{\alpha,\beta}, \\
\{a_\alpha, a_\beta\} &= \{a_\alpha^\dagger, a_\beta^\dagger\} = 0.
\end{aligned} \tag{11.18}$$

The curly brackets denote "anti"-commutation, i.e. $\{A,B\} = AB + BA$. One consequence of the anti-commutation relations is that any time two creation operators or two destruction operators with the same index are next to one another the result is zero, $a_\alpha a_\alpha = a_\alpha^\dagger a_\alpha^\dagger = 0$. This enforces the Pauli exclusion principle. Similar to the Bose example, the vacuum is annihilated by the destruction operator.

$$a_\alpha|0\rangle = 0. \tag{11.19}$$

The commutation rules written above assumed the states $\alpha, \beta, \cdots$ were orthogonal. It is more complicated if the indices refer to states from different bases. The first basis will be labeled with roman letters $i, j, k \cdots$ and the second will be with greek, $\alpha, \beta, \gamma \cdots$. First, as seen in Chapter 9, one can see that under a basis transformation creation and destruction operators transform like single-partcle states,

$$\begin{aligned} \langle \alpha | &= \sum_k \langle \alpha | k \rangle \langle k |, \\ \langle 0 | a_\alpha &= \langle 0 | \sum_k a_k \langle \alpha | k \rangle, \\ a_\alpha &= \sum_k \langle \alpha | k \rangle a_k. \end{aligned} \tag{11.20}$$

The anti-commutation algebra becomes

$$\{a_\alpha, a_j^\dagger\} = \sum_k \langle \alpha | k \rangle \{a_k, a_j^\dagger\} = \langle \alpha | j \rangle. \tag{11.21}$$

We are now able to calculate arbitrary matrix elements for N particle states,

$$\begin{aligned} \langle i, j, k, \cdots | \alpha, \beta, \gamma, \cdots \rangle &= \langle 0 | (a_i a_j a_k \cdots)(a_\gamma^\dagger a_\beta^\dagger a_\alpha^\dagger \cdots) | 0 \rangle \\ &= \sum_{\text{perm.s of } i,j,\cdots} (-1)^{n_p} \langle i | \alpha \rangle \langle j | \beta \rangle \langle k | \gamma \rangle \cdots \end{aligned} \tag{11.22}$$

As an example we consider two-particle matrix elements,

$$\langle i, j | \alpha, \beta \rangle = \langle i | \alpha \rangle \langle j | \beta \rangle - \langle j | \alpha \rangle \langle i | \beta \rangle. \tag{11.23}$$

If i and j refered to positions x and y, the result would read

$$\langle x, y | \alpha, \beta \rangle = \phi_\alpha(x)\phi_\beta(y) - \phi_\alpha(y)\phi_\beta(x), \tag{11.24}$$

which signifies that the two particles can not be at the same position. Note that the overlap of an N-particle matrix element would yield $N!$ terms. Thus, it seems little has been gained using field operators rather than writing symmetrized/anti-symmetrized wave functions which also have $N!$ terms. However, when writing the bra and ket with all permutations, both the bra and ket have $N!$ terms, with the extra $N!$ being cancelled by the $1/\sqrt{N!}$s in the normalization of the wave functions. This rapidly becomes onerous.

It should also be noted that all the same results are valid for bosons, except that the $(-1)^{n_p}$ factors disappear.

11.6 Fermionic Field Operators

The field operators $\Psi_s^\dagger(x)$ and $\Psi_s(x)$ create and destroy a particle at position x with spin s. They obey the anti-commutation relations,

$$\{\Psi_s(\vec{x}), \Psi_{s'}^\dagger(\vec{y})\} = \delta^3(\vec{x} - \vec{y})\delta_{ss'}. \tag{11.25}$$

Just like the non-fermionic operators defined in Sec. 9.2, they can be defined in terms of creation and destruction operators for momentum states,

$$\Psi_s^\dagger(\vec{r}) = \frac{1}{\sqrt{V}} \sum_{\vec{k}} e^{-i\vec{k}\cdot\vec{r}} a_s^\dagger(\vec{k}). \tag{11.26}$$

The field operators commute with other operators as

$$\{\Psi_s(\vec{x}), a_\alpha^\dagger\} = \phi_{\alpha,s}(\vec{x}), \tag{11.27}$$

where $\phi_{\alpha,s}(\vec{x}) = \langle \vec{x}, s|\alpha\rangle$ is the single particle wave function of a particle of spin s in the state α. Thus, the field operators are no different than other creation and destruction operators, except that their dimension is length$^{-3/2}$ and their anti commutation relations are expressed in terms of Dirac deltas instead of Kronecker deltas.

Considering a state where the ortho-normal single-particle states are $\alpha, \beta \cdots$, the expectation of the density operator is

$$\rho(\vec{x}) = \sum_s \Psi_s^\dagger(\vec{x})\Psi_s(\vec{x}). \tag{11.28}$$

The expectation of the density operator in the state $|\alpha, \beta, \cdots ...\rangle$ is

$$\begin{aligned}\langle \alpha, \beta, \gamma, \cdots |\Psi_s^\dagger(\vec{x})\Psi_s(\vec{x})|\alpha, \beta, \gamma, \cdots\rangle &= \langle 0|(a_\alpha a_\beta a_\gamma \cdots)\Psi_s^\dagger(\vec{x})\Psi_s(\vec{x})(\cdots a_\gamma^\dagger a_\beta^\dagger a_\alpha^\dagger)|0\rangle \\ &= \sum_{\kappa\in\alpha,\beta\cdots} \phi_{\kappa,s}^*(\vec{x})\phi_{\kappa,s}(\vec{x}).\end{aligned} \tag{11.29}$$

Then density operator can also appear in a transition element. Consider the bra and ket to differ in one label, $\gamma \to \gamma'$,

$$\begin{aligned}\sum_s \langle \alpha, \beta, \gamma', \cdots |\Psi_s^\dagger(\vec{x})\Psi_s(\vec{x})|\alpha, \beta, \gamma, \cdots\rangle &= \langle 0|(a_\alpha a_\beta a_{\gamma'} \cdots)\Psi^\dagger(\vec{x})\Psi(\vec{x})(\cdots a_\gamma^\dagger a_\beta^\dagger a_\alpha^\dagger)|0\rangle \\ &= \sum_s \phi_{\gamma',s}^*(\vec{x})\phi_{\gamma,s}(\vec{x}).\end{aligned} \tag{11.30}$$

This simple result followed because it was assumed that the final single-particle states in the bra, were used to describe the ket. Thus, the N-particle state was specified by listing N single-particle labels. In more realistic calculations, one might study a transition between states where many of the wave functions were continually altered as particles were added. For instance, in a many-electron atom, if one adds the N^{th} particle it affects the structure of the previous $N-1$ particles. This assumption allows one to write the state as a single product of creation operators acting on the vacuum. A product state is produced by operating with a single string of orthogonal creation operators on the vacuum. More generally, a state might be comprised from a linear combination of product states. Going beyond simple product states requires sophisticated many-body treatments, which is beyond the scope of this class.

Example 11.5: Correlations in a Fermi Gas
Here, we calculate correlations in a Fermi gas due to the Fermi nature of the particles. Consider a one-dimensional gas of fermions moving in a large region of length L where the density corresponds to a Fermi momentum k_f.

(a) Find the correlation function, relating the ratio of the probability of finding two particles separated by r to the probability of finding the two particles at locations far removed from each other.

Solution: To do this problem, we first write the two-particle probability,

$$\rho_2(x_1, x_2) = \langle\phi|\Psi^\dagger(x_1)\Psi^\dagger(x_2)\Psi(x_2)\Psi(x_1)|\phi\rangle.$$
$$|\phi\rangle = \prod_{k<k_f} a_k^\dagger|0\rangle.$$

Commuting the field operators $\Psi(x_2)$ and $\Psi(x_1)$ toward the ket yields

$$\Psi(x_2)\Psi(x_1)\left(\prod_{k<k_f} a_k^\dagger\right)|0\rangle = \sum_{k_1,k_2<k_f} \frac{e^{ik_2x_2}}{\sqrt{L}}\frac{e^{ik_1x_1}}{\sqrt{L}} a_{k_2}a_{k_1}\left(\prod_{k<k_f} a_k^\dagger\right)|0\rangle$$
$$= \sum_{k_1,k_2<k_f} (\pm)(-)^{n_1+n_2}\frac{e^{ik_2x_2}}{\sqrt{L}}\frac{e^{ik_1x_1}}{\sqrt{L}}\left(\prod_{k<k_f,k\neq k_1,k\neq k_2} a_k^\dagger\right)|0\rangle.$$

Here, the sign $(\pm)$ depends on whether $a(k_1)a(k_2)$ appear in the same/opposite order as $a^\dagger(k_1)a^\dagger(k_2)$ in the product of creation operators. One can find a similar expression for $\langle\phi|\Psi^\dagger(x_1)\Psi^\dagger(x_2)|\phi\rangle$,

$$\langle 0|\left(\prod_{k'<k_f} a_{k'}\right)\Psi^\dagger(x_1)\Psi^\dagger(x_2)$$
$$= \langle 0|\sum_{k_1',k_2'<k_f}\left(\prod_{k'<k_f,k'\neq k_1',k\neq k_2'} a_{k'}\right)(\pm)(-)^{n_1'+n_2'}\frac{e^{-ik_2'x_2}}{\sqrt{L}}\frac{e^{-ik_1'x_1}}{\sqrt{L}}.$$

The overlap of the two states is zero unless k_1 and k_2 from the ket are matched with k_1' and k_2' in the bra. However, this can happen with either $k_1 = k_1', k_2 = k_2'$, or with $k_1 = k_2', k_2 = k_1'$. For the first ordering, the two phases exactly cancel. The signs of $(-1)^{n\cdots}$ are also identical and cancel. For the second ordering, the product of phases yields $e^{i(k_1-k_2)(x_1-x_2)}$ results, plus there is an extra sign due to the permutation of the creation operators, $a_{k_1'} \leftrightarrow a_{k_2'}$. This then yields

$$\langle\phi|\Psi^\dagger(x_1)\Psi^\dagger(x_2)\Psi(x_2)\Psi(x_1)|\phi\rangle = \sum_{k_1,k_2<k_f}\frac{1}{L^2}\left(1 - e^{i(k_1-k_2)(x_1-x_2)}\right).$$

Because for every pair k_1, k_2, the same pair appears again as k_2, k_1. Thus, only the real part of the phase factor $e^{i(k_1-k_2)(x_1-x_2)}$ enters, and the matrix element is indeed Hermitian.

Taking the ratio of the two-particle density to the square of the one particle densities gives the correlation function.

$$g(r \equiv x_2 - x_1) \equiv \frac{P_2(x_1, x_2)}{n^2} = \frac{\sum_{k_1,k_2<k_f} \left(1 - e^{i(k_1-k_2)(x_1-x_2)}\right)}{\sum_{k_1,k_2<k_f}}.$$

The sums can be changed into integrals,

$$g(r) = 1 - |I|^2,$$
$$I = \frac{\int_{-k_f<k<k_f} dk e^{ikr}}{\int_{-k_f<k<k_f} dk} = \frac{\sin(k_f r)}{k_f r}.$$

The two-particle density is thus

$$\rho_2(x_1, x_2) = n^2 \left\{1 - \left[\frac{\sin(k_f r)}{k_f r}\right]^2\right\}, \quad r \equiv |x_1 - x_2|.$$

The correlation is displayed in Fig. 11.2.

(b) The particles now feel a mutual interaction,

$$V(r) = \alpha\Theta(a - r),$$

where r is again the relative distance between two particles. Find the exchange contribution to the energy per unit length due to the interaction in first-order perturbation theory.

Solution: In first-order perturbation theory the correction to the energy is

$$\langle H\rangle = \frac{1}{2}\int dx_1 dx_2 \langle\phi|\Psi^\dagger(x_1)\Psi^\dagger(x_2)\Psi(x_2)\Psi(x_1)|\phi\rangle V(x_1 - x_2)$$
$$= L\frac{1}{2}\int dr \langle\phi|\Psi^\dagger(0)\Psi^\dagger(r)\Psi(r)\Psi(0)|\phi\rangle V(r),$$

where we have used the fact that the two-particle density depends only on $r = x_1 - x_2$.

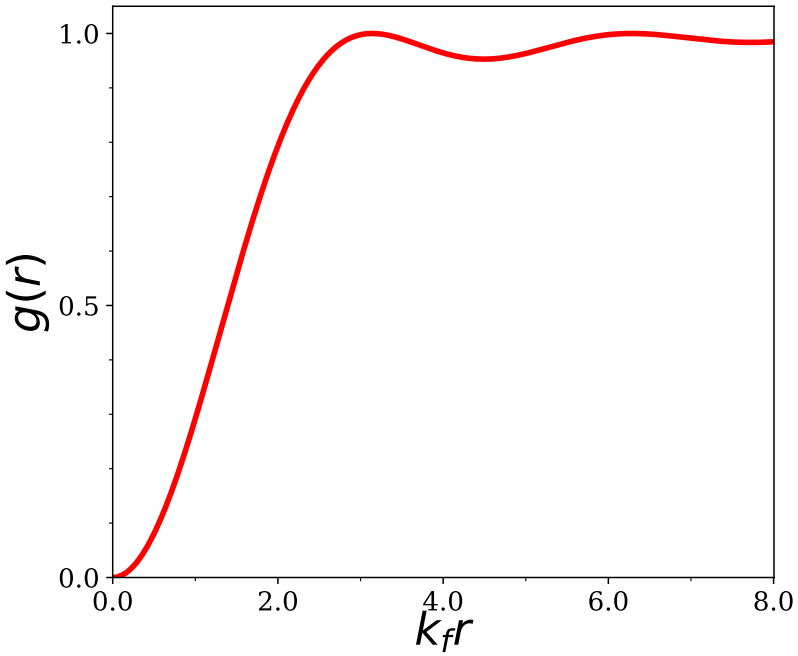


Figure 11.2: The electron-electron correlation (for electrons of one sign) due to anti-symmetrization in a one-dimensional system, as derived in Example 11.5.

By using the results of (a) one can find the energy per length,

$$\begin{aligned}
\frac{\langle H\rangle}{L} &= \frac{1}{2}n^2\int_{-\infty}^{\infty} dr\, g(r)V(r) \\
&= \frac{\alpha n^2}{2}\int_{-a}^{a} dr\, \left\{1-\left[\frac{\sin(k_f r)}{k_f r}\right]^2\right\} \\
&= \alpha n^2\left\{a+\frac{1}{k_f^2 r}\sin^2(k_f r)\Big|_0^a - \int_0^a dr\frac{2\sin(k_f r)\cos(k_f r)}{k_f r}\right\} \\
&= \alpha n^2\left\{a+\frac{\sin^2(k_f a)}{k_f^2 a}-\frac{1}{k_f}\mathrm{Si}(2k_f a)\right\} \\
&= \alpha n^2\left\{a+\frac{\sin^2(\pi n a)}{\pi^2 n^2 a}-\frac{1}{\pi n}\mathrm{Si}(2\pi n a)\right\},
\end{aligned}$$

where Si is the sine-integral function, and n is the number of particles per unit length. Note that the first term describes the potential energy one would expect from particles with a uniform density interacting through the potential $V(x_1 - x_2)$. This is called the direct term, while the remainder is referred to as the exchange term.

11.7 Exercises

1. Consider a two-dimensional zero-temperature non-relativistic gas of identical spin-up fermions whose mass is m. They are confined in a two-dimensional box of dimensions, L_x and L_y. The quantum numbers characterizing the single-particle eigenstates are n_x and n_y. The box is divided in half along the x axis. The eigenstates with odd n_x now disappear, while there are now two solutions (one for each half of the box) for each even value of n_x. Assuming the size of the box is large compared to the inverse Fermi momentum, find the penalty, expressed as an energy per unit length ($\Delta E/2L_y$), for dividing the box. Express your answer in terms of the Fermi momentum and the mass.
2. In the interior of a neutron star, the neutron-to-proton ratio is very high. This results despite the fact that the proton mass is 1.3 MeV higher than the neutron mass. This occurs because protons must balanced by an equal number of electrons. Furthermore, protons and neutrons may be interchanged via the reaction,

   $$p + e \leftrightarrow n$$

   (We have neglected the neutrinos in this reaction because they are free to leave the star due to their massless nature.)

   The masses of the particles are:

   $$m_p c^2 = 938.27\,\text{MeV},\ \ m_n c^2 = 939.57\,\text{MeV},\ m_e c^2 = 0.511\,\text{MeV}$$

   For the problems below, you may assume the protons and neutrons are non-relativistic, $E = mc^2 + (\hbar ck)^2/(2mc^2)$, but the electrons must be treated as relativistic particles, $E = [(\hbar ck)^2 + (mc^2)^2]^{1/2}$. It is useful to remember that $\hbar c = 197.326$ MeV·fm.

   (a) If the baryon(neutrons or protons) density equals n_B, express the corresponding constraint involving the Fermi momenta of the protons and neutrons.
   (b) From the constraint that the system is electrically neutral, describe how the proton's Fermi momentum is related to the electron's Fermi momentum.
   (c) Describe how minimizing the overall energy results in a constraint involving the three Fermi momenta.
   (d) For 4 cases, $n_B = 0.0001, 0.001, 0.1, 1.0$ baryons per cubic Fermi, solve the expressions above and plot the neutron/proton ratio as a function of n_B. Solving the three equations may involve finding roots numerically.

3. Calculate the magnetic susceptibility of a two-dimensional electron gas.
4. Consider the wave function of three identical particles in Eq. (11.15) where the three single-particle wave functions are orthonormalized.

   (a) Show that

   $$\frac{1}{3!}\int d^3r_1 d^3r_2 d^3r_3\ |\phi_{\alpha,\beta,\gamma}(\vec{r}_1,\vec{r}_2,\vec{r}_3)|^2 = 1 \tag{11.31}$$

   if the three indices are different, $\alpha \neq \beta, \alpha \neq \gamma, \beta \neq \gamma$, and that the overlap is zero if any two of the indices are the same.

(b) Show that $\phi_{\alpha,\beta,\gamma}(\vec{r}_1, \vec{r}_2, \vec{r}_3) = 0$ if any two positions are the same.

5. Correlation/anticorrelation in a quantum gas: Consider a uniform gas of non-interacting spin-half particles in the ground state. The wave function may be written

$$|\phi\rangle = \prod_{\alpha, |k_\alpha| < k_f} a_\alpha^\dagger |0\rangle,$$

where the product includes all states α with momentum $k_\alpha < k_f$ and spin s_α. The density-density correlation function is defined as

$$g_{s_1,s_2}(\vec{x_2} - \vec{x}_1) \equiv \frac{\langle\phi|\Psi_{s_1}^\dagger(\vec{x}_1)\Psi_{s_2}^\dagger(\vec{x}_2)\Psi_{s_2}(\vec{x}_2)\Psi_{s_1}(\vec{x}_1)|\phi\rangle}{\langle\phi|\Psi_{s_1}^\dagger(\vec{x}_1)\Psi_{s_1}(\vec{x}_1)|\phi\rangle\langle\phi|\Psi_{s_2}^\dagger(\vec{x}_2)\Psi_{s_2}(\vec{x}_2)|\phi\rangle}$$

This function expresses the correlation of two particles with spin s_1 and s_2 being separated by $\vec{x}_2 - \vec{x_1}$. It is defined in such a way that it is unity if the probability of seeing two particles at $\vec{x}_1$ and $\vec{x}_2$ is the product of the probabilities of observing each particle independently.

(a) Show that the density-density correlation function can be written as

$$g_{s_1,s_2}(\vec{x_2} - \vec{x}_1) = 1 - \delta_{s_1 s_2} \frac{\sum_{k_\alpha, k_\beta} e^{i(\vec{k}_\alpha - \vec{k}_\beta)\cdot(\vec{x}_2 - \vec{x}_1)}}{\sum_{k_\alpha, k_\beta}},$$

where the sums are over all momentum states with $k_\alpha < k_f$ and $k_\beta < k_f$.

(b) By changing the sum over states to a three-dimensional integral over $\vec{k}$, find an analytic expression for the density-density correlation function in terms of the Fermi momentum k_f.

6. Again, consider a non-interacting quantum gas of particles of mass m, with the ground state being expressed as

$$|\phi\rangle = \prod_{\alpha, |k_\alpha| < k_f} a_\alpha^\dagger |0\rangle,$$

where the product includes all states α with momentum $k_\alpha < k_f$. For this problem, ignore the spin indices. Consider an interaction between the particles of the form,

$$H_{\text{int}} = \frac{1}{2} \int d^3r_1 d^3r_2 V(\vec{r}_1 - \vec{r}_2)\Psi^\dagger(\vec{r}_1)\Psi^\dagger(\vec{r}_2)\Psi(\vec{r}_2)\Psi(\vec{r}_1)$$
$$V(\vec{r}_1 - \vec{r}_2) = \beta\delta(\vec{r}_2 - \vec{r}_1)$$

Find the first-order perturbative correction for the energy, $\langle\phi|H_{\text{int}}|\phi\rangle$, for particles in the gas in terms of k_f, β and m.

12 Non-Perturbative Approaches for Many-Fermion Systems

In the previous chapter, interactions between fermions were mainly ignored. A one-body interaction is typically introduced via a spatial potential $V(\vec{r})$. Even if this is a potential that depends on the relative position of two particles, $V(\vec{r}_2 - \vec{r}_2)$, one can work in center of mass coordinates and effectively reduce the problem to a two-body problem to a one-body problem (at least in non-relativistic physics). Bound state and scattering observables can then all be calculated with theories that involve only the relative coordinate, $\vec{r} = \vec{r}_1 - \vec{r}_2$. As soon as more than two particles are present, the problem no longer reduces to a one-body problem. In this chapter we focus mainly on such problems where there is an external potential and ignore interaction terms between constituents. This allows one to assume that the many-particle wave functions are products, or anti-symmetrized products, of single particle wave functions. More generally, interaction terms simultaneously modify the behavior of more than one particle at a time and the true solutions cannot be expressed as products. Such treatments of more realistic interactions are typically under the moniker of "many-body theory", which is largely beyond the scope of this course. Nonetheless, the one-body treatments, although somewhat crude, provide a great deal of insight and in some cases can be remarkably accurate. One example is the Hartree-Fock approximation, which is the topic of the Sec. 12.3.

12.1 Interacting Fermi Systems – Two and Four-Point Interactions

The simplest sort of interaction is a "two-point" interaction.

$$H = \sum_{\alpha,\beta} V_{\alpha\beta} a^\dagger_\alpha a_\beta. \tag{12.1}$$

A two-point interaction has one creation and one destruction operator in each term. The interaction alters the momentum of a single particle, which makes it a one-body interaction. The particles behave independently. To demonstrate what is meant by "independent", consider the product state

$$|\phi(t=0)\rangle = a^\dagger_\delta \cdots a^\dagger_\beta a^\dagger_\alpha |0\rangle. \tag{12.2}$$

Imagine one were to make a linear transformation of the creation operators using a unitary matrix U,

$$b^\dagger_\alpha = U_{\alpha\beta} a^\dagger_\beta. \tag{12.3}$$

These new operators would also satisfy the anti-commutation laws,

$$\begin{aligned} \{b_\alpha, b^\dagger_\beta\} &= U^\dagger_{\gamma\alpha} U_{\beta\delta} \{a_\gamma, a^\dagger_\delta\} \\ &= \{a_\alpha, a^\dagger_\beta\} = \delta_{\alpha\beta}. \end{aligned} \tag{12.4}$$

Thus, a unitary matrix acting on the space of creation operators results in a new set of operators that still obey the anti-commutation laws. Further, if that unitary matrix is chosen to be the same

matrix that diagonalizes $V_{\alpha\beta}$,

$$\begin{aligned} H &= \sum_{\alpha\beta} V_{\alpha\beta} U^\dagger_{\gamma\alpha} U_{\beta\delta} b^\dagger_\gamma b_\delta \\ &= (U^\dagger V U)_{\alpha\beta} b^\dagger_\alpha b_\beta. \end{aligned} \tag{12.5}$$

Thus, if there is a two-body Hamiltonian, one can transform to a basis where the Hamiltonian is diagonalized, i.e.

$$H = \sum_\alpha E_\alpha b^\dagger_\alpha b_\alpha. \tag{12.6}$$

Eigenstates of the $n-body$ problem can then be expressed as a product state,

$$|\alpha\beta...\gamma\rangle = \prod_{\alpha\beta\cdots\gamma} b^\dagger_\alpha b^\dagger_\beta \cdots b^\dagger_\gamma |0\rangle, \tag{12.7}$$

and the eigenenergies are $E = E_\alpha + E_\beta + \cdots + E_\gamma$.

This is not the case if the Hamiltonian has terms such as

$$V_{\alpha\beta\gamma\delta} a^\dagger_\alpha a^\dagger_\beta a_\gamma a_\delta. \tag{12.8}$$

These terms take a pair of particles in one-body states γ and δ and put them into states α and β. Scattering does precisely this. If one writes a two-body potential with field operators

$$V = \frac{1}{2} \int d^3r d^3r'\, V(\vec{r} - \vec{r}') \Psi^\dagger(\vec{r}) \Psi^\dagger(\vec{r}') \Psi(\vec{r}') \Psi(\vec{r}), \tag{12.9}$$

four-body terms are inevitable.

However if the the potential is from an external one-body potential,

$$V = \int d^3r\, V(r) \Psi^\dagger(\vec{r}) \Psi(\vec{r}), \tag{12.10}$$

only two-body terms appear, and one can find eigenstates that are products of creation and destruction operators. Unfortunately, physics is not usually so simple. If all interactions were of the two-point type, all problems could be solved by diagonalizing a matrix.

An interaction between particles separated by $\vec{r} = \vec{r}_1 - \vec{r}_2$ is written classically as

$$H_{\text{int}} = \frac{1}{2} \int d^3r_1 d^3r_2 \rho(\vec{r}_1) \rho(\vec{r}_2) V(\vec{r}_1 - \vec{r}_2). \tag{12.11}$$

The factor of $1/2$ corrects for double counting because the integral is without a $r_1 < r_2$ qualifier.

Written in terms of field operators, this interaction becomes

$$H_{\text{int}} = \frac{1}{2} \int d^3r_1 d^3r_2 \Psi^\dagger(\vec{r}_1) \Psi^\dagger(\vec{r}_2) V(\vec{r}_1 - \vec{r}_2) \Psi(\vec{r}_2) \Psi(\vec{r}_1). \tag{12.12}$$

For the moment, we will omit spin indices and assume only particles of a specific spin are involved. One peculiar aspect of the Hamiltonian is that the two $\Psi^\dagger$s are on the left while the two

Ψs are on the right, whereas the product of densities would suggest a product of $\Psi^\dagger\Psi$s. However, such a product would result in an interaction energy for a one-particle state, as the particle would interact with itself.

Four-point interactions (or three-point interactions, e.g. the electromagnetic interaction $\vec{J}\cdot\vec{A}$) are infinitely more complicated than two-point interactions. For one, if one wishes to calculate the evolution of $a^\dagger_\alpha(t)$, one needs to commute the Hamiltonian with $a^\dagger_\alpha$. However, if $H_{\rm int}$ is a product of four creation and destruction operators, commuting $H_{\rm int}$ with a_α or $a^\dagger_\alpha$ results in a product of three operators. This makes it impossible to consider the problem as a simple mixing of the creation and destruction operators. The eigenstates of a Hamiltonian with three-point or four-point terms is not a product state, but is instead a complicated linear combination of product states. Such problems are rarely solvable, forcing one to resort to approximations. This is what is meant by the many-body problem, a moniker which encompasses most of the effort in theoretical physics. Examples of such approximation methods are perturbation theory or the Hartree-Fock approximation.

12.2 The Thomas-Fermi Approximation

A very crude way to find the density of fermions in an external well is the Thomas-Fermi approximation. This approximation is semiclassical and assumes that the local density is a function of the local potential only,

$$n(\vec{r}) = \frac{2s+1}{6\pi^2}k_f(\vec{r})^3,\ \ \frac{\hbar^2 k_f(\vec{r})^2}{2m} + V(\vec{r}) = \mu_f. \tag{12.13}$$

The density would vanish outide the region where $V(\vec{r}) < \mu_f$. Given the potential and the chemical potential μ_f, it is straight forward to find the density. For a finite system with a fixed particle number, one would have to adjust μ_f to get the correct total number of particles. The Thomas-Fermi approximation is crude, and gives rather non-sensical results for atoms due to the small number of electrons in atoms. It is less unreasonable for estimating the electron density near surfaces. As a general rule it is valid if the systems are large and the potential changes slowly.

Example 12.1: The Thomas Fermi Approximation for the Harmonic Oscillator
Calculate the Fermi energy (as measured from the bottom of the harmonic oscillator) for N electrons in a one-dimensional harmonic oscillator characterized by angular frequency ω.
Solution: The density of electrons for a potential $V(x)$ in one dimension is:

$$\begin{aligned} n(x) &= (2s+1)\frac{1}{2\pi}\int_{-k_f}^{k_f} dk \\ &= \frac{2}{\pi}k_f(x), \\ &= \frac{2}{\hbar\pi}\sqrt{2m(\mu_f - V(x))}. \end{aligned}$$

For the harmonic oscillator, this gives

$$\begin{aligned} N &= \int_{-x_{\max}}^{x_{\max}} dx\, n(x) \\ &= \frac{4\sqrt{2m\mu_f}}{\hbar\pi} \int_0^{x_{\max}} dx \sqrt{1-\alpha^2 x^2}, \\ \alpha &= \sqrt{\frac{m\omega^2}{2\mu_f}}, \quad x_{\max} = \sqrt{\frac{2\mu_f}{m\omega^2}}. \end{aligned}$$

Substituting $u = \alpha x$,

$$\begin{aligned} N &= \frac{8\mu_f}{\hbar\omega\pi} \int_0^1 du \sqrt{1-u^2} \\ &= \frac{2\mu_f}{\hbar\omega}. \end{aligned}$$

If one were to set $\mu_f = (n+1/2)\hbar\omega$, the true ground state energy of the harmonic oscillator, one would find $N = 2n+1$. Whereas if one filled all the levels, each with a spin-up and a spin-down fermion, the number of particles would be $2n$. This is a good approximation for large N, but is poor for a small number of particles. The conditions under which this method becomes accurate is similar to the conditions discussed for the WKB approximation in Sec. 6.1.

12.3 The Hartree-Fock Approximation

The Hartree-Fock approximation yields an expression, that when solved gives an approximation to the ground state of a many-fermion system. The approximation is variational in nature. But rather than varying a few parameters to modify the wave function, the value of the wave-function at each point is treated as a variational parameter. In a variational approximation one assumes some form for the variational wave function. In this case, one assumes that the solution is in some sort of product state,

$$|\phi\rangle = a_\delta^\dagger \cdots a_\beta^\dagger a_\alpha^\dagger |0\rangle. \tag{12.14}$$

The many-particle wave function, $\langle \vec{r}_1 \cdots \vec{r}_n | a_n^\dagger \cdots a_1^\dagger |0\rangle$, will not be expressed as a product of single-particle wave functions. But that is only a matter of anti-symmetrization. Given one of the simple product terms, the other terms of the many-particle wave-function can be generated by permuting the labels of the first term.

To derive the Hartree Fock equations, one writes down the expectation of the Hamiltonian,

$$\begin{aligned} \langle\phi|H|\phi\rangle = \langle\phi| \int d^3r \Psi^\dagger(\vec{r}) \left(-\frac{\hbar^2}{2m}\nabla^2\right) \Psi(\vec{r}) + \int d^3r U(r) \Psi^\dagger(\vec{r})\Psi(\vec{r})|\phi\rangle \\ + \langle\phi| \frac{1}{2} \int d^3r_1 d^3r_2 V(\vec{r}_2 - \vec{r}_1) \Psi^\dagger(\vec{r}_1)\Psi^\dagger(\vec{r}_2)\Psi(\vec{r}_2)\Psi(\vec{r}_1)|\phi\rangle, \end{aligned} \tag{12.15}$$

where the spin indices have been omitted temporarily.

One can now express the energy in terms of the wave functions of the occupied states α.

$$\langle\phi|\boldsymbol{H}|\phi\rangle = \sum_{\alpha}\int d^3r\,\phi^*_{\alpha}(\vec{r})\left(-\frac{\hbar^2}{2m}\boldsymbol{\nabla^2}+\boldsymbol{U}(\vec{r})\right)\phi_{\alpha}(\vec{r}) \tag{12.16}$$
$$+\frac{1}{2}\sum_{\alpha,\alpha'}\int d^3r d^3r'\,\boldsymbol{V}(\vec{r}-\vec{r}')\left\{\phi^*_{\alpha}(\vec{r})\phi^*_{\alpha'}(\vec{r}')\phi_{\alpha'}(\vec{r}')\phi_{\alpha}(\vec{r})\pm\phi^*_{\alpha}(\vec{r})\phi^*_{\alpha'}(\vec{r}')\phi_{\alpha}(\vec{r}')\phi_{\alpha'}(\vec{r})\right\}.$$

The second term in the two-body interaction, where the $\pm$ refers to bosons/fermions, is known as the exchange term, with the name coming from the exchange of the α and α' indices.

The next step in the variational procedure is to minimize $\langle\phi|\boldsymbol{H}|\phi\rangle$ with respect to changes in the wave functions subject to the constraint that each wave function is properly normalized. One accounts for the constraint by multiplying the constraint by a Lagrange multiplier $\boldsymbol{\lambda}_{\gamma}$ then adding it to the function one wishes to minimize,

$$\frac{\boldsymbol{\delta}}{\boldsymbol{\delta}\phi_{\gamma}(\vec{r})}\left\{\langle\phi|\boldsymbol{H}|\phi\rangle-\boldsymbol{\lambda}_{\gamma}\int d^3r\phi^*_{\gamma}(\vec{r})\phi_{\gamma}(\vec{r})\right\}=\boldsymbol{0}. \tag{12.17}$$

More correctly, one would vary both the real and imaginary parts of ϕ_{γ} which is equivalent to varying either ϕ_{γ} or ϕ^*_{γ}. The resulting expression is

$$\boldsymbol{\lambda}_{\gamma}\phi_{\gamma}(\vec{r}) = \left(-\frac{\hbar^2}{2m}\boldsymbol{\nabla^2}+\boldsymbol{U}(\vec{r})\right)\phi_{\gamma}(\vec{r}) \tag{12.18}$$
$$+\sum_{\alpha'}\int d^3r'\boldsymbol{V}(\vec{r}-\vec{r}')\left\{\phi^*_{\alpha'}(\vec{r}')\phi_{\alpha'}(\vec{r}')\phi_{\gamma}(\vec{r})\pm\phi^*_{\alpha'}(\vec{r}')\phi_{\gamma}(\vec{r}')\phi_{\alpha'}(\vec{r})\right\}.$$

The two terms involving $\boldsymbol{V}(\vec{r}-\vec{r}')$ are known as the Hartree and Fock terms respectively. The Hartree term looks like a potential felt by the particles due to the presence of the other particles. Thus, the Hartree equation can be solved by self-consistently finding the solutions to the effective potential

$$\boldsymbol{V}_{\mathbf{Hartree}}(\vec{r})=\sum_{\alpha'}\int d^3r'\boldsymbol{V}(\vec{r}-\vec{r}')\phi^*_{\alpha'}(\vec{r}')\phi_{\alpha'}(\vec{r}') \tag{12.19}$$

The Hartree equations are often solved iteratively. One guesses at the wave functions, finds the Hartree-potential, solves the Schrödinger equation for the single-particle wave functions, then iterates the procedure until the wave functions converge. One issue with the Hartree approximation is that a particle also feels the potential due to itself unless one restricts the sum over α' to ignore the term $\alpha=\gamma$ in Eq. (12.18). If one excludes this term, then the potential is different for each state, which then leads to the solutions ϕ_{γ} not being orthogonal.

The Fock term presents a different challenge because $\phi_{\gamma}(\vec{r})$ cannot be factored out of the differential equation. The Fock term is non-local in that $\phi_{\gamma}(\vec{r}')$ appears in its place.

The Lagrange multiplier $\boldsymbol{\lambda}_{\gamma}$ plays the role of a single-particle energy. However, one should remember that because the interaction energy between two states is contained in the single-particle energy for both states, the total energy is not the sum of the $\boldsymbol{\lambda}_{\gamma}$s. The Hartree-Fock approximation is especially useful for estimating density distributions or even changes in energies (e.g. separation energies), but is not particularly accurate at calculating total binding energies.

Finally, we return to the problem of including the spin indices. When spin indices are included the Fock term disappears unless the two states are the same species and the same spin. Hence a factor $\delta_{s,s'}$ accompanies the Fock term.

12.4 Hartree-Fock for Atoms

The principal difficulty in handling multi-electron atoms comes from the fact that the electrons interact with one another. If it were not for the mutual interactions between electrons, one could treat the electrons independently. Of course, this would yield horribly unphysical results. For instance, one could place an infinite number of electrons into the the hydrogen atom, where in reality no more than two electrons can be bound to hydrogen.

The Hartree approximation is the first method one might consider for treating such atoms. One then solves the Schrödinger equation with the potential,

$$V_i(\vec{r}) = -\frac{Ze^2}{r} + \sum_{j\neq i} \int d^3r' |\phi_j(\vec{r}')|^2 \frac{e^2}{|\vec{r}-\vec{r}'|} \tag{12.20}$$

If there are n electrons, one must solve n coupled equations.

One can go one step further and solve the Hartree-Fock equations,

$$\left(-\frac{\hbar^2\nabla^2}{2m} - \frac{Ze^2}{r}\right)\phi_i(\vec{r}) + \sum_j \int d^3r' \frac{e^2}{|\vec{r}-\vec{r}'|}\phi_j^*(\vec{r}')\left[\phi_j(\vec{r}')\phi_i(\vec{r}) - \delta_{s,s'}\phi_j(\vec{r})\phi_i(\vec{r}')\right] = \lambda_i\phi_i(\vec{r}) \tag{12.21}$$

to find the wave functions. One can then find the energy by calculating $\langle\phi|\boldsymbol{H}|\phi\rangle$. Again, this requires solving coupled differential equations, but the Fock term also makes the equations non-local which brings along an added computational difficulty.

The difference of the energy found by solving the Hartree-Fock equations and the Hartree equations is the exchange energy. Because the Hartree-Fock wave functions force the electrons to stay away from one another due to anti-symmetrization, the repulsive Coulomb interaction between electrons is weakened which means that the exchange energy is negative, affecting the solution like an attractive contribution to the potential.

The Hartree-Fock equations were derived with the assumption that the solution is a product state. As that is a variational assumption, the true ground state energies are about 1 eV lower than Hartree-Fock solutions. This difference is referred to as "correlation" energy. It can be calculated perturbatively by considering the mixing of one-particle-one-hole and two-particle-two-hole states into the wave functions. This is done perturbatively by considering the n-electron Hartree-Fock ground state as the ground state. Then one considers excited states as being those n-particle n-hole states formed by considering n of the single-particle solutions of the Hartree-Fock ground state to be replaced by n single-particle solutions from the set of single-particle states left unfilled in the Hartree-Fock ground state.

12.5 The Periodic Table

Single-electron wave functions are labeled by n, ℓ, s, j and m_j. In an atom with only one electron, the energy depends principally on n as states with same n but different $\ell = 0, 1, \cdots n-1$ are degenerate aside from the spin-orbit interaction. Of course, s always equals $1/2$ and j can be either $\ell + 1/2$ or $\ell - 1/2$.

The existence of the other electrons destroys the "accidental" degeneracy of the hydrogen atom and allows states with different ℓ and the same n to have significantly different energy. This is due to the screening of the positive charge. The "accidental" degeneracy allowed states with fewer nodes in the radial wave function but larger angular momentum to have the same energy as states with smaller angular momentum but more nodes in the radial wave function. By screening the charge, an advantage is created for states that have a relatively greater probability of being near the origin. Because radial wave functions behave as r^ℓ near the origin, a state with a lower ℓ but the same n will have lower energy due to screening. This difference can be large enough at times to allow a state to move lower than states of higher n but higher angular momentum. For instance, in some cases the $4s$ states can move below the $3d$ states, and similarly the $4d$ and $5s$ shells compete as well. The $5s$ shell is always well below the $4f$ shell. When shells compete, e.g. the $4s$ and $3d$ electrons, the choice of orbitals is non trivial. In these cases the configuration can vary from one element to the next, and in fact, solutions might contain a mixture of configurations. Chemical properties are determined largely by the outermost electrons. When shells are filled, the elements are less reactive. The rare gases (also know as inert or noble gases) all have filled p shells with the exception of Helium.

The set of orbitals with a specific n and ℓ is known as a "shell". The degeneracy of a shell is $4\ell + 2$. Electronic "configurations" are labeled by the shells and the filling, e.g.,

$$(1s)^2(2s)^2(2p)^3,$$

with the superscript labeling the number of electrons in the shell.

12.6 Configuration Splitting and Hund's Rules

Different configurations are generally split by a few electron volts. There are $4\ell + 2$ single-particle orbitals in a shell, and the number of ways to arrange several electrons among these levels can be rather large. For multi-electron atoms one first considers the single-particle orbitals with a given radial quantum number n and orbital quantum number ℓ (of course $s = 1/2$). If all the levels of a given shell are filled, the net angular momentum of that shell is zero. But if the shell is not either completely filled or completely empty, one might be able to configure the electrons into numerous states. The energies of these configurations depends on the total (of the particles in the semi-filled shell) spin angular momentum S, the total orbital angular momentum L, and the total momentum J. Ignoring the spin-orbit interaction, the total angular momentum $\vec{L}$ and total spin $\vec{S}$ of the electrons commute with the Hamiltonian. The $(2L+1)(2S+1)$ states of an LS multiplet are then further split by the spin-orbit interaction. Thus, the spin-orbit interaction invalidates M_L and M_S as good quantum numbers, and replaces them with J and M_J.

Hund's rules determine which LSJ combination has the lowest energy:

1. The largest $\boldsymbol{S}$ is most favorable as it has the most symmetric spin wave functions, therefore totally anti-symmetric spatial wave functions. This minimizes the probability that electrons are close to one another, and therefore minimizes their repulsive interaction.
2. For a given $\boldsymbol{S}$, the $\boldsymbol{LS}$ multiplet with the largest $\boldsymbol{L}$ has the lowest energy. This can be understood by considering two electrons rotating about a nucleus. If $\boldsymbol{L}$ is large, both electrons are moving about with essentially the same orbit as $\boldsymbol{m_1}$ and $\boldsymbol{m_2}$ can be large and of the same sign. As such electrons spend less time crossing one another's path they miminize their repulsive interaction.
3. Different $\boldsymbol{LS}$ multiplets are typically split by tenths of electron volts. A given $\boldsymbol{LS}$ multiplet is again split by the spin-orbit interaction but only by a few hundredths or a few thousandths of an electron volt. The spin-orbit interaction gives preference to states with the minimum $\boldsymbol{J}$ when the shells are less than half filled and behaves in the opposite way when the shells are more than half filled.

Configurations are labeled by $\boldsymbol{S}$, $\boldsymbol{L}$ and $\boldsymbol{J}$, which should not be confused with the ℓ of the single-particle orbitals. The notation is

$$^{2S+1}\boldsymbol{L_J}.$$

For instance, the state $^3\boldsymbol{D_3}$ would have the spins coupled to $\boldsymbol{S=1}$, $\boldsymbol{L=2}$ and $\boldsymbol{J=3}$. Note the upper-case angular momentum labels $\boldsymbol{S, P, D, F}\cdots$.

It should be remembered that considering only ways to arrange particles in a given shell is an approximation. There is no reason that the true ground state shouldn't have some mixture of states from different shells.

12.7 Constructing States According to Total $\vec{L}$ and Total $\vec{S}$

The overall wave function for a many-electron state described by $\boldsymbol{L}$, $\boldsymbol{S}$ and $\boldsymbol{J}$ can be written as the sum over products of angular and spin wave functions.

$$|\boldsymbol{L, S, J, M_J}\rangle = \sum_{\boldsymbol{M_L, M_S}} \langle \boldsymbol{J, M_J}, |\boldsymbol{L, M_L, S, M_S}\rangle |\boldsymbol{L, M_L}\rangle |\boldsymbol{S, M_S}\rangle \tag{12.22}$$

The angular/spin wave functions $|\boldsymbol{L, M_L}\rangle|\boldsymbol{S, M_S}\rangle$ must first be written in terms of products of singular particle wave functions summed over with the help of Clebsch-Gordan technology.

As an example we consider a two-electron state with orbital/spin wave functions in a shell of

angular momentum ℓ.

$$\begin{aligned}\Psi_{L,S,J,M_J}(\Omega,\Omega';m_s,m_s') &= \langle \Omega,\Omega';m_s,m_s'|L,S,J,M_J\rangle \\ &= \sum_{M_L,M_S}\langle J,M_J|L,M_L,S,M_S\rangle\langle\Omega,\Omega'|L,M_L\rangle\langle m_s,m_s'|S,M_S\rangle \\ &= \sum_{M_L,M_S}\langle J,M_J|L,M_L,S,M_S\rangle \\ &\cdot \sum_{m_\ell,m_\ell'}\langle L,M_L|\ell,m_\ell,\ell,m_\ell'\rangle Y_{\ell,m_\ell}(\Omega)Y_{\ell,m_\ell'}(\Omega') \\ &\cdot \langle S,M_S|1/2,m_s,1/2,m_s'\rangle \\ &= \sum_{M_L,M_s,m_\ell,m_\ell'}\langle J,M_J|L,M_L,S,M_S\rangle\langle L,M_L|\ell,m_\ell,\ell,m_\ell'\rangle \\ &\qquad \langle S,M_S|1/2,m_s,1/2,m_s'\rangle Y_{\ell,m_\ell}(\Omega)Y_{\ell,m_\ell'}(\Omega').\end{aligned} \tag{12.23}$$

The expression would be even more complicated if there were three electrons in the shell.

Example 12.2: Electrons in Carbon
Here we construct the ${}^1S_0$ state in Carbon. The two electrons in the p shell of Carbon can be in any number of configurations. Using the expression above,

$$\begin{aligned}&\Psi_{L=0,S=0,J=0,M_J=0}(\Omega,\Omega',m_s,m_s') \\ &= \sum_{m_\ell,m_\ell'}\langle J,M_J|L,M_L,S,M_S\rangle\langle L=0,M_L=0|\ell=1,m_\ell,\ell=1,m_\ell'\rangle \\ &\qquad \cdot\langle 1/2,1/2,S=0,M_S=0|m_s,m_s'\rangle Y_{\ell=1,m_\ell}(\Omega)Y_{\ell=1,m_\ell'}(\Omega') \\ &= \sum_{m_\ell,m_\ell'}\langle L=0,M_L=0|\ell=1,m_\ell,\ell=1,m_\ell'\rangle Y_{\ell=1,m_\ell}(\Omega)Y_{\ell=1,m_\ell'}(\Omega') \\ &\qquad \cdot\frac{1}{\sqrt{2}}\left(\delta_{m_s,1/2}\delta_{m_s',-1/2}-\delta_{m_s,-1/2}\delta_{m_s',1/2}\right) \\ &= \frac{1}{\sqrt{6}}\left\{Y_{1,1}(\Omega)Y_{1,-1}(\Omega')+Y_{1,-1}(\Omega)Y_{1,1}(\Omega')-Y_{1,0}(\Omega)Y_{1,0}(\Omega')\right\} \\ &\qquad \cdot\left(\delta_{m_s,1/2}\delta_{m_s',-1/2}-\delta_{m_s,-1/2}\delta_{m_s',1/2}\right)\end{aligned} \tag{12.24}$$

Note that the wave function is symmetric with respect to interchange of Ω with Ω' and anti-symmetric with respect to interchange of m_s and m_s', making the overall wave function anti-symmetric.

12.8 Permutation Symmetry

When coupling two particles together the permutation symmetry goes as $(-1)^L$ for the spatial part, while the interchange of spins is symmetric/anti-symmetric for $S = 1/0$. For this reason one can notice that for two electrons in the (2p) shell (carbon) the orbital states must be $L = 0$

or $L = 2$ if the the spin state is $S = 0$ and must be $L = 1$ if $S = 1$. Thus, the possible states in carbon are ${}^1S_0$, ${}^1D_2$, ${}^3P_0$, ${}^3P_1$ and ${}^3P_2$.

Coupling a higher number of particles together can lead to "mixed" symmetries. In this case the angular wave functions may be neither symmetric or anti-symmetric, while the spin wave functions might be mixed as well. However, the overall wave function needs to be anti-symmetric. For example, the flavor and spin wave functions of the three quarks that constitute a proton or neutron are in states of mixed symmetry.

12.9 Zeeman Effect and the Landé g Factor

A particle in a magnetic field feels the interaction,

$$H_{\rm mag} = -\frac{eB}{2mc}(L_z + 2S_z), \tag{12.25}$$

where the magnetic field is assumed to point in the z direction. We wish to calculate the change in energy for an atom in a state of good J, M_J due to the interaction.

$$\begin{aligned} \Delta E_{LSJM_J} &= -\frac{eB}{2mc}\langle L, S, J, M_J | L_z + 2S_z | L, S, J, M_J\rangle \\ &= -\frac{eB}{2mc}\left(\hbar M_J + \langle L, S, J, M_J | S_z | L, S, J, M_J\rangle\right). \end{aligned} \tag{12.26}$$

The challenge in calculating the splitting comes from finding $\langle S_z\rangle$. We expect this to be proportional to M_J as it is the only label available. The Wigner Eckart theorem and Clebsch-Gordan technology come to the rescue.

$$\frac{\langle L, S, J, M_J | S_z | L, S, J, M_J\rangle}{\langle L, S, J, M_J | J_z | L, S, J, M_J\rangle} = \frac{\langle L, S, J || S || L, S, J\rangle}{\langle L, S, J || J || L, S, J\rangle}. \tag{12.27}$$

This follows by applying the Wigner Eckart theorem to both $\langle S_z\rangle$ and $\langle J_z\rangle$ and noticing that the same Clebsch-Gordan coefficients appear in both terms.

Thus, if we can find the ratio of the reduced matrix elements we will have fulfilled our mission. To do this, we first step aside to perform a proof. Consider a vector operator $\vec{A}$, which will be chosen to be $\vec{S}$ further below.

$$\begin{aligned} \langle JM | \vec{A}\cdot\vec{J} | JM\rangle &= \sum_{-1\le m'\le 1} \langle JM | A_{m'} J_{-m'} | JM\rangle \\ &= \sum_{-1\le m'\le 1, -J\le M''\le J} \langle JM | A_{m'} | JM''\rangle\langle JM'' | J_{-m'} | JM\rangle, \end{aligned} \tag{12.28}$$

Note that the inserted states were only those within the same JM multiplet. This is valid because J_m does not mix different multplets. Now, by applying the Wigner Eckart theorem, one

can write the matrix element $\langle \vec{A} \cdot \vec{J} \rangle$ as

$$\begin{aligned} \langle JM|\vec{A}\cdot\vec{J}|JM\rangle &= f(J,M)\langle J||A||J\rangle\langle J||J||J\rangle \\ f(J,M) &= \sum_{-1\le m'\le 1, -J\le M''\le J} \\ &\quad \cdot \langle 1,J,J,M|m',M''\rangle\langle 1,J,J,M|-m',M''\rangle \frac{1}{2J+1} \end{aligned} \tag{12.29}$$

Because f is determined solely by Clebsch-Gordan coefficients, one can see that

$$\frac{\langle JM|\vec{A}\cdot\vec{J}|JM\rangle}{\langle JM|\vec{J}\cdot\vec{J}|JM\rangle} = \frac{\langle J||A||J\rangle}{\langle J||J||J\rangle} \tag{12.30}$$

Now that our proof is finished we can see that

$$\begin{aligned} \frac{\langle L,S,J||S||L,S,J\rangle}{\langle L,S,J||J||L,S,J\rangle} &= \frac{\langle L,S,J,M_J|\vec{J}\cdot\vec{S}|L,S,J,M_J\rangle}{\langle L,S,J,M_J|\vec{J}\cdot\vec{J}|L,S,J,M_J\rangle} \\ &= \frac{J(J+1)+S(S+1)-L(L+1)}{2J(J+1)}, \end{aligned} \tag{12.31}$$

where we have taken advantage of the fact that

$$\begin{aligned} \vec{J}\cdot\vec{S} &= \frac{1}{2}\left[-(\vec{J}-\vec{S})^2+\vec{J}^2+\vec{S}^2\right] \\ &= \frac{\hbar^2}{2}[J(J+1)+S(S+1)-L(L+1)]. \end{aligned} \tag{12.32}$$

Finally, we are able to insert our result for the ratio of the reduced matrix elements into our expression for the splitting to get

$$\begin{aligned} \Delta E &= -g\frac{e\hbar B}{2mc}M_J \\ g &= 1+\frac{J(J+1)+S(S+1)-L(L+1)}{2J(J+1)}. \end{aligned} \tag{12.33}$$

The factor g is known as the Landé g factor, with the g referring to *gyration*.

In some undergraduate modern physics books the Landé g factor is derived from simple geometric arguments. If one considers the vector J precessing about the z axis and the vector S precessing about $\vec{J}$, one would expect that the expectation of $\vec{S}$ for a given orientation of $\vec{J}$ would be

$$\langle \vec{S} \rangle = \vec{J}\frac{\vec{S}\cdot\vec{J}}{|J|^2}, \tag{12.34}$$

because as $\vec{S}$ precesses about $\vec{J}$ the component of $\vec{S}$ perpendicular to $\vec{J}$ averages to zero. Finally, averaging over the direction of $\vec{J}$ gives the result we expected.

12.10 Molecules and the Adiabatic Approximation

The force felt between atoms largely derives from the distortion of the electronic wave functions due to the proximity of a second potential. Because atoms move slowly with respect to the electronic motion, electronic wave functions largely adjust in such a way as to remain in the ground state for a given atomic separation. If N atoms are centered at positions $\vec{R}_1 \cdots \vec{R}_N$, one can calculate the net electronic binding energy $E(\vec{R}_1 \cdots \vec{R}_N)$ for the given positions of the atoms. One then solves a Schrödinger equation for the atomic positions, $\Psi(\vec{R}_1 \cdots \vec{R}_N)$, using the net electronic binding energy as the potential. This is referred to as the adiabatic or Born-Oppenheimer approximation. Examples where the adiabatic approximation were applied included a particle in a slowly expanding box. If a box were slowly expanded or contracted, a particle would remain in the ground state of the box. Of course, this implies that the energy of the particle is not kept constant. For the example of an expanding box, the ground state energy falls as the volume expands, and the lost energy appears in the kinetic energy of the piston. In fact, that energy is equal to the work, PdV, done by the expansion.

In the case of atoms, the gain or loss of the electronic energy appears as a loss or gain of the ion's kinetic energy. The electronic energy, calculated as a function of the ion's separation R, serves as a potential for the atoms. In a previous exercise, 6.8, one was asked to find the energy felt by two electrons in harmonic oscillator potentials situated far apart, such that they felt a dipole-dipole interaction. Calculating the correction to the energy, it was found that the energy went as $1/R^6$. This energy would then serve as a potential energy for understanding the behavior of two atoms if each well was given a mass and one were to solve a Schroödinger equation using this energy as a potential. However, in this case this only gives the potential a larger distances. Different approaches would be necessary to estimate the electronic energy for two harmonic oscillators, or two atoms, brought into closer contact.

Example 12.3: Interacting with Distant Ion
For this example we consider an electron in a well interacting with a distant positive ion, where we assume the states available to the electron are described as those of a harmonic oscillator with frequency ω. This differs from the exercise 6.8 in that the distant particle is not another charged particle in a well, which would be overall neutral, but carries a charge Ze. The ion is placed a distance R from the atom. The perturbative potential between the ion and the atom is

$$\begin{aligned} H_{\text{int}} &= Ze^2 \left\{ \frac{1}{R} - \int d^3r \rho(\vec{r}) \frac{1}{|\vec{R} - \vec{r}|} \right\} \\ &\approx -Ze^2 \int d^3r \rho(\vec{r}) \frac{z}{R^2} \\ &= -\frac{Ze^2}{R^2} z_{\text{op}}. \end{aligned}$$

One can then calculate the energy to second order perturbation theory,

$$\begin{aligned}\Delta E^{(2)} &= -\frac{Z^2e^4}{R^4}\frac{|\langle n_z=1|z_{\text{op}}|0\rangle|^2}{\hbar\omega} \\ &= -\frac{Z^2e^4}{2R^4m\omega^2}.\end{aligned}$$

This energy serves as the potential between the atom and the ion. It shows that the well is attracted toward the distant ion. Note that if one were to consider an induced dipole moment proportion to the electric field, multiplied by the electric field which falls off as $1/R^2$, one would have expected the $1/R^4$ behavior. This can be considered as an interaction of a charge with an induced dipole moment, as opposed to the exercise 6.8, which was a dipole-dipole interaction, or more accurately an induced-dipole-induced-dipole interaction.

Example 12.4: ^2H positive ion
Here, we consider the $^+H_2$ ion, where two protons are separated by a distance R, with the addition of a single electron. In the adiabatic approximation the binding energy of the electron, $\epsilon(R)$, serves as the potential between the protons when combined with the proton-proton interaction. At large distances the binding energy is one Rydberg, while one would expect an increase in the magnitude of the binding energy as R is shortened. At very short distances, one expects the interaction to become repulsive when one penetrates the electronic cloud and the interaction is dominated by the Coulomb interaction between the protons.
Calculating the binding energy as a function of R can be crudely accomplished with a variational calculation assuming a trivial form of the wave function,

$$\psi(\vec{r}) = C_\pm\left[\psi_A(\vec{r}) \pm \psi_B(\vec{r})\right],$$

where ψ_A and ψ_B are the bound state wave functions of an electron to each of the protons,

$$\begin{aligned}\psi_A(\vec{r}) &= (\pi a_0^3)^{-1/2}e^{-|\vec{r}-\vec{R}_A|/a_0}, \\ \psi_B(\vec{r}) &= (\pi a_0^3)^{-1/2}e^{-|\vec{r}-\vec{R}_B|/a_0}.\end{aligned}$$

The factor $C_\pm$ is merely a normalization constant,

$$\begin{aligned}C_\pm &= \frac{1}{\sqrt{2\pm 2S(R)}} \\ S(R) &= \int d^3r\psi_A(\vec{r})\psi_B\vec{r}) \\ &= \left(1+\frac{R}{a_0}+\frac{R^2}{3a_0^2}\right)e^{-R/a_0}.\end{aligned}$$

This can be considered as a variational calculation with zero variational parameters.
This integral $S(R)$ is most easily calculated in elliptic coordinates, where the three compo-

nents of $\vec{r}$ are replaced by

$$\begin{aligned} u &\equiv \frac{|\vec{r}-\vec{R}_A| + |\vec{r}-\vec{R}_B|}{R}, \\ v &\equiv \frac{|\vec{r}-\vec{R}_A| - |\vec{r}-\vec{R}_B|}{R}, \\ \phi &\equiv \arctan y/x. \end{aligned}$$

Thus, u is the scaled sum of the distances to the two protons, v is the scaled difference of the two distances and ϕ is the usual azimuthal angle with respect to the axis defined by the two protons. Some Jacobian manipulations would reveal,

$$\int d^3r f(\vec{r}) = \int_1^\infty du \int_{-1}^1 dv \int_0^{2\pi} d\phi \frac{R^3}{8}(u^2-v^2) f(\vec{r}).$$

Once one has made a transformation into these coordinates the integral becomes rather trivial because

$$\psi_A(\vec{r})\psi_B(\vec{r}) = (\pi a_0^3)^{-1} e^{-uR/a_0}$$

If you have not worked with elliptic coordinates before, you can understand the name "elliptical" because the set of points with a fixed sum of distances, $|\vec{r}-\vec{R}_A| + |\vec{r}-\vec{R}_B|$, defines an ellipse.
One is now in the position to calculate the expectation of the energy,

$$\begin{aligned} \langle H \rangle_\pm = \epsilon_\pm(R) &= \frac{\langle A|H|A\rangle + \langle B|H|B\rangle \pm 2\langle A|H|B\rangle}{2 \pm 2S} \\ &= \frac{\langle A|H|A\rangle \pm \langle A|H|B\rangle}{1 \pm S}, \end{aligned}$$

where

$$\begin{aligned} \langle A|H|A\rangle &= \epsilon_1 + \frac{e^2}{R} - \int \psi_A^2(\vec{r}) \frac{e^2}{|\vec{r}-\vec{R}_B|} d^3r \\ &= \epsilon_1 + \frac{e^2}{R}\left(1 + \frac{R}{a_0}\right) e^{-2R/a_0}, \end{aligned}$$

and

$$\begin{aligned} \langle A|H|B\rangle &= \left(\epsilon_1 + \frac{e^2}{R}\right) S - \int d^3r \psi_A(\vec{r})\psi_B(\vec{r}) \frac{e^2}{|\vec{r}-\vec{R}_B|} \\ &= \left(\epsilon_1 + \frac{e^2}{R}\right) S - \frac{e^2}{a_0}\left(1 + \frac{R}{a_0}\right) e^{-R/a_0}. \end{aligned}$$

The term ϵ_1 is merely the electronic binding energy of one electron with a proton, -1.0 Rydbergs. The integrals were calculated with the help of the tranformation into elliptic coordinates.

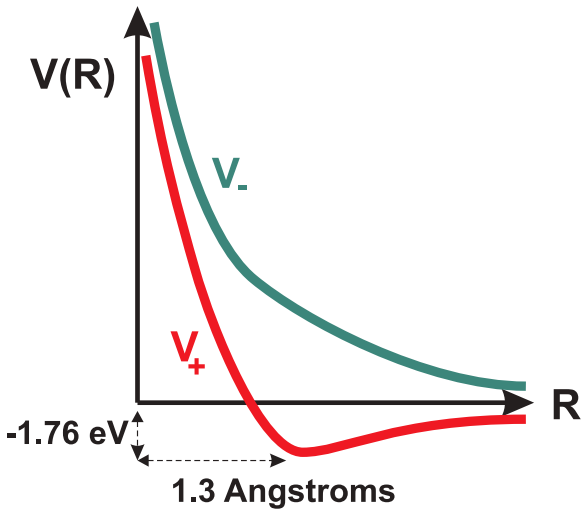


Figure 12.1: Potential for $^{+}\boldsymbol{H_2}$ ion from Example 12.4.

The potential between the atoms is effectively

$$\boldsymbol{V_{\pm}(R) = \langle H \rangle_{\pm} - \epsilon_1.}$$

When plotted against $\boldsymbol{R}$ the resulting curves are shown in Fig. 12.1.
In the adiabatic approximation, where the mass of the proton is considered to be very large, the binding energy should simply be the minimum of the potential. In reality the binding energy is deeper by an entire eV. However, one could improve the calculation by considering a variational calculation.
The difference between $\boldsymbol{V_+}$ and $\boldsymbol{V_-}$ comes from the fact that the parities of the two wave functions were positive and negative respectfully. The positive-parity solution has lower energy because the probability density is larger in the region between the protons where the interaction with the protons is maximized, and because the negative-parity solution has an extra node, which increases the kinetic energy.
This result for the potential should be contrasted with the previous example. In that case the potential at large distances was found to fall as $\boldsymbol{1/R^4}$ using perturbation theory. Even though that calculation was based on the harmonic oscillator, one would still expect that for a hydrogen atom the perturbative result to scale as the square of the electric field and behave as $\boldsymbol{1/R^4}$. In this calculation the potential fell exponentially at large $\boldsymbol{R}$, so it rather poorly predicts the behavior for large distances. This failure has nothing to do with the adiabatic nature of the approximation, but rather with the poor choice of wave functions in the variational approximation.

12.11 The Hydrogen Molecule and Pairing

The neutral hydrogen molecule has two electrons. One can pursue a variational calculation as performed in the previous problem; only in this instance the variational wave function is a two-electron wave function. When we discussed multi-electron atoms, Hund's first rule stated

that the prime criteria for multi-electon states to minimize the energy is to form states of largest $\boldsymbol{S}$ so that the spatial wave function will be maximally anti-symmetrized thus minimizing the electron-electron repulsive interaction. However, in the case of the $\boldsymbol{H_2}$ molecule, the driving determination is that both electrons should be in the even-parity ψ_+ state, which through the Pauli exclusion principle requires that the electrons be in a spin singlet with $\boldsymbol{S = 0}$.

When two electrons form a spin-singlet they are called a pair. Much of the systematics of molecular binding can be understood by considering pairing. Pairing that involves $\boldsymbol{p}$ states is somewhat more complicated. Of the three $\boldsymbol{p}$ states, the $\boldsymbol{m_\ell = 0}$ state extends along the $\boldsymbol{z}$ axis and thus more strongly samples the attractive interaction with the other proton. Pairing between the $\boldsymbol{m_\ell = \pm 1}$ states can also take place but is usually weaker. The two pairings are refered to as $\boldsymbol{\sigma}$ and $\boldsymbol{\pi}$ bondings. The $\boldsymbol{\pi}$ label refers to the fact that the transverse pairing allows an angular momentum about the molecular axis.

For atoms where $\boldsymbol{p}$ and $\boldsymbol{s}$ states are nearly degenerate, pairing with neighboring atoms can be especially strong. In these cases the wave function can be a linear combination of an $\boldsymbol{s}$ state and a $\boldsymbol{p}$ state with $\boldsymbol{m_\ell = 0}$. Separately, both states have equal weight on the $\boldsymbol{z}$-positive and $\boldsymbol{z}$-negative sides of the atom. However, linear combination of an even-parity and an odd-parity state will result in a wave function with higher probability on one side. Given the presence of a neighboring atom, this can lead to an energetically favorable situation. By taking a linear combination of such a "hybrid" state with a hybrid state using the opposing atom as it's center, one can make states with overall good reflection symmetry about the half-way point between the atoms.

12.12 Superconductivity

Pairing of electrons is responsible for the phenomena known as superconductivity. The connection between forming pairs and reducing the conductivity to zero is actually rather subtle. The formed pairs, which can have zero net momentum, can coalesce and move as a coherent unit through the material. Here, we concentrate on the pairing mechanism and offer only the following rather heuristic discussion of how this leads to superconductivity.

One could imagine resistance, or dissipation, coming from three sources:

1. As we will see, a gap energy will be associated with the pairs. In order for the moving pairs to de-excite one must excite the pairs by at least the gap energy. If one imagines a current moving through a circular loop at a temperature near zero, there is not enough thermal energy available to break a pair. Because the pairs are bosons they can condense into a single zero-momentum state. The pairs join together into a condensate in such a way that not only is there an energy penalty to break a pair, but there is a penalty to remove a pair from the condensate. Thus, at small temperature, it is extremely difficult to de-excite the condensate by breaking pairs or removing pairs from the condensate.
2. One could imagine a drag force acting on the condensate due to moving the condensate through the background of non-condensate particles. If the condensate is treated as a macroscopic object, the drag force would then behave as

$$\boldsymbol{F_{\rm drag} \propto u^2, \;\; P = Fv \propto u^3,} \tag{12.35}$$

where u is the condensate velocity. Because the power scales as u^3 and because the resistance is defined as

$$\begin{aligned} P = I^2 R, R = P/I^2, \\ I \propto u,\ P \propto u^3 \\ R = \frac{P}{I^2} \propto u \to 0. \end{aligned} \tag{12.36}$$

Thus, drag can be neglected with respect to its contribution to the resistance. Because the coalesced pairs move as a unit, they can carry a large current with a very small velocity and thus vanishingly small drag.

3. The pairs condense into a coherent unit, and one could imagine exciting a pair, while leaving it bound. However, because of attractive inter-pair interactions, the pairs in the condensate also require energy to be dislodged. This energy tends to exceed that required to break a pair.

The best way to dissipate movement of the condensate is to break a pair. This requires an amount of energy 2Δ, where Δ is known as the gap energy. Only those electrons with energies greater than the gap can contribute to the dissipation. The conductivity then behaves as $e^{-2\Delta/T}$

The phenomena of superconductivity has much in common with superfluidity. In both cases, it is rather easy to motivate why one forms a condensate, but explaining why the condensate moves without dissipation is subtle. Both an electron pair and a ^{4}He atom are bosonic, but formed of an even number of fermions. Superconductivity might occur in any Fermi system, and is thought to occur due to the pairing of protons or neutrons in the crusts of neutron stars. When nuclear matter is compressed to extremely high density, several times higher than the density of matter in nucleus, and becomes a Fermi gas of quarks, one also expects pairing of the quarks.

12.13 Cooper Pairs

Our goal in this section will not be to understand the correlated structure of the pairs, but only to show that the existence of pairs lowers the energy compared to a Fermi gas. In order for electrons to pair there must exist an attractive interaction between two electrons with momenta near the Fermi surface. The idea of two electrons interacting attractively is most peculiar because electrons have the same charge. The excuse for assuming an attractive interaction is that the electrons interact with one another via the lattice. An electron might exchange momentum with an ion, which then translates that momentum into a sound wave, or phonon, which travels through the lattice. The phonon then interacts with the other electron through its interaction with some other ion. Intuitively, one might expect such an interaction to provide an effective polarizability that reduces the electric repulsion, but does not reverse it. However, electrons are free to move and screen the repulsive electric interaction, as in Debye screening. This allows the attractive interaction through phonons to overwhelm the electromagnetic interaction.

For our purposes, we follow the work of Bardeen, Cooper and Schraefer (BCS) and assume a simplified interaction between electrons,

$$\langle \vec{k}_1 \vec{k}_2 | \mathcal{V} | \vec{k}'_1 \vec{k}'_2 \rangle = \begin{cases} -\frac{v_0}{V} \delta_{k_1+k_2,k'_1+k'_2}, & k_f < k_1, k_2, k'_1, k'_2 < k_a \\ 0, & \text{otherwise} \end{cases} \tag{12.37}$$

where v_0 sets the scale of the interaction, V is the volume, and k_a is close to k_f. Thus, the model assumes the attractive interaction is confined to particles within the neighborhood of the Fermi surface.

Because total momentum is a good quantum number, eigenstates can have the form,

$$|\vec{K}\text{ paired}\rangle = {\sum_{\vec{k}'}}' a_{\vec{k}'}(\vec{K})|\vec{k}'_1 = \vec{K}/2 + \vec{k}', \vec{k}'_2 = \vec{K}/2 - \vec{k}'\rangle, \tag{12.38}$$

where the primed sum is over all relative momenta $\vec{k}' \equiv (\vec{k}'_1 - \vec{k}'_2)/2$ such that both $\vec{k}'_1$ and $\vec{k}'_2$ are inside the region between k_f and k_a.

Solving for the eigenstates,

$$\begin{aligned} H|\vec{K}\text{ paired}\rangle &= E|\vec{K}\text{ paired}\rangle, \\ -\frac{v_0}{V}{\sum_{\vec{k}'}}' a_{\vec{k}'}(\vec{K}) &= (E - \epsilon_{\vec{k}_1} - \epsilon_{\vec{k}_2}) a_{\vec{k}}(\vec{K}) \end{aligned} \tag{12.39}$$

Dividing both sides by $(E - \epsilon_{\vec{k}_2})$, and summing both sides over $\vec{k}$,

$${\sum_{\vec{k}}}' a_{\vec{k}}(\vec{K}) - \frac{v_0}{V}{\sum_{\vec{k}}}' \frac{1}{E - \epsilon_{\vec{k}_1} - \epsilon_{\vec{k}_2}} {\sum_{\vec{k}'}}' a_{\vec{k}'}(\vec{K}) = 0. \tag{12.40}$$

Now, one can see how the simplifying assumption that the matrix element $v_{k,i'}$ was independent of k and k' in the subspace simplifies the problem. It allows the sum over amplitudes to be canceled from both sides of the equation above and result in,

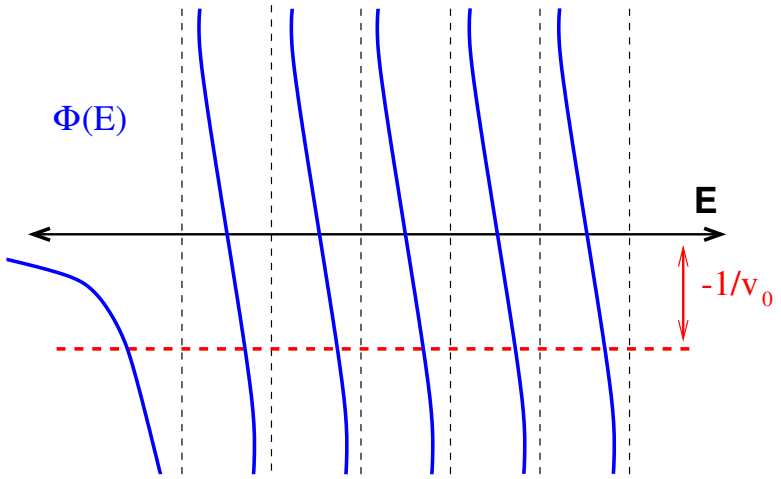
$$\begin{aligned} -\frac{1}{v_0} &= \frac{1}{V}\sum_{\vec{k}'} \frac{1}{E - \epsilon_{\vec{k}_1} - \epsilon_{\vec{k}_2}} \\ &= \frac{1}{V}{\sum_{\vec{k}}}' \frac{1}{E - \epsilon_{\vec{K}+\vec{k}} - \epsilon_{\vec{K}-\vec{k}}} \\ &= \frac{1}{V}\sum_{\vec{k}} \frac{1}{E - \hbar^2K^2/(4m) - \hbar^2k^2/m}, \end{aligned} \tag{12.41}$$

where the last step involved writing the energy as a center-of-mass energy plus an energy-of-relative-motion.

One can solve for E in the equation above graphically. To illustrate the solutions, we consider the function

$$\Phi_K(E) \equiv \frac{1}{V}{\sum_{\vec{k}}}' \frac{1}{E - \hbar^2K^2/(4m) - \hbar^2k^2/m}, \tag{12.42}$$

and graph the function to see for what energies, E, the function $\Phi(E)$ equals $-1/v_0$.



Every time E passes an energy, $\hbar^2K^2/(4m) + \hbar^2k^2/m$, Φ changes from $-\infty$ to $+\infty$. Thus, for every value of $\vec{k}$ there exists a solution with $E < \hbar^2K^2/(4m) + \hbar^2k^2/m$ where $\Phi = -1/v_0$. The vertical dashed lines in the figure represent the energies $\hbar^2K^2/(4m) + \hbar^2k^2/m$, which would be the eigenenergies if v_0 were to be zero. The intersections of the blue curves with the horizontal dashed line represent the solutions.

If $\vec{K} = 0$, the first value of $\vec{k}$ that enters the primed sum is $|\vec{k}| = k_f$. Thus, there exists a solution with energy $E < 2\epsilon_f$, even though the solution was formed from momentum states above k_f. The paired is energetically favorable compared to being a momentum state at the Fermi surface. Note that as v_0 is increased the solution for the energy becomes lower.

12.14 The Gap

The expression for Φ in Eq. (12.42) can be integrated analytically for the case with $K = 0$ if one approximates the sum over states as

$$\sum_{\vec{k}}{}' \to \rho(\epsilon_f) \int_{\epsilon_f}^{\epsilon_a} d\epsilon_k. \tag{12.43}$$

This amounts to assuming the density of states, $\rho(\epsilon_k)$, is constant in the region of integration. The expression for Φ then becomes

$$\begin{aligned} \Phi(E) &= \rho(\epsilon_f) \int_{\epsilon_f}^{\epsilon_a} dE_k \frac{1}{E - 2\epsilon_k} \\ &= -\frac{\rho(2\epsilon_f)}{2} \log\left|\frac{2\epsilon_a - E}{2\epsilon_f - E}\right|, \end{aligned} \tag{12.44}$$

where

$$\rho(\epsilon_f) = \frac{mk_f}{2\pi^2\hbar^3}. \tag{12.45}$$

Solving for E,

$$E = \frac{-2\epsilon_a + 2\epsilon_f e^{2/(\rho v_0)}}{e^{2/(\rho v_0)} - 1}. \tag{12.46}$$

The difference between this energy and $2\epsilon_f$ is known as the gap energy, Δ.

$$\begin{aligned}\Delta &\equiv \frac{2\epsilon_f - E}{2} \\ &= \frac{\epsilon_a - \epsilon_f}{e^{2/[\rho(\epsilon_f)v_0]} - 1}.\end{aligned} \tag{12.47}$$

One can do a similar calculation for $K \neq 0$, but the binding energy would be smaller because the interaction links fewer pairs as both electrons in the pair, with momenta $k_\pm = K/2 \pm k$, must lie within the shell in momentum space.

One can see that the energy required to break the pair is 2Δ, because in order to break the pair one must move the electrons outside the Fermi sea, that is, increase their energy beyond $2\epsilon_f$.

12.15 Landau Levels and the Integer Quantum Hall Effect

Consider particles of mass m and charge e moving in a two-dimensional world, the x-y plane. A magnetic field is present, $\vec{B} = B\hat{z}$. The Hamiltonian for such particles is

$$H = \left(\vec{P} - e\vec{A}/c\right)^2/(2m). \tag{12.48}$$

If one chooses as a vector potential,

$$\vec{A} = Bx\hat{y}, \tag{12.49}$$

one obtains the desired magnetic field. Because there is no y dependence in the Hamiltonian, solutions can be chosen as eigenstates of the operator P_y with eigenvalue p_y. One can then rewrite the Hamiltonian,

$$H = \frac{P_x^2}{2m} + \frac{(p_y - eBx/c)^2}{2m}, \tag{12.50}$$

where p_y is simply a number and P_x is an operator. Because p_y is a number, one can consider the second term as an offset harmonic oscillator potential,

$$\begin{aligned}H &= \frac{P_x^2}{2m} + \frac{1}{2}m\omega^2\left(x - \frac{p_y c}{eB}\right)^2, \\ \omega &\equiv \frac{eB}{mc}.\end{aligned} \tag{12.51}$$

Thus, the Hamiltonian looks like a one-dimensional Harmonic oscillator with a frequency equal to the Larmor frequency. The harmonic oscillator is centered at $x = p_y c/(eB)$.

Note that the eigenenergies, $(n + 1/2)\hbar\omega$, do not depend on p_y. Thus, there are many solutions with different p_y that have identical energies. The energy levels are referred to as Landau levels.

Each level has a degeneracy equal to the number of values of p_y for which there exist solutions. Using the fact that the $y-$dependence behaves as $e^{ip_y y/\hbar}$ and applying the boundary conditions that the wave function vanishes at $y = 0$ and $y = L_y$, one finds the same values of p_y as one would have for particles in a box of size L_y. The density of such states is then

$$\frac{dN}{dp_y} = \frac{L_y}{2\pi\hbar}, \tag{12.52}$$

where L_y is the size of the sample in the y direction. The limits on p_y are determined by the x dimension. Because p_y determines the center of the center of the Harmonic oscillator, $x = p_y c/(eB)$, and because the centers must be between $x = 0$ and $x = L_x$,

$$0 < \frac{p_y c}{eB} < L_x, \tag{12.53}$$
$$0 < p_y < p_{y,\max} = \frac{eBL_x}{c}.$$

The number of such states is then

$$N = \frac{dN}{dp_y} p_{y,\max} = \frac{eBL_x L_y}{2\pi\hbar c}. \tag{12.54}$$

Assume the sample has a number of conduction electrons per unit area n. If the number of conduction electrons exactly fits an integer number of Landau levels, the levels will be exactly filled and the conductance changes. Thus, varying the magnetic field one sees changes in the conductance when

$$nL_x L_y = m\frac{eBL_x L_y}{2\pi\hbar c}, \quad \text{m is an integer}, \tag{12.55}$$
$$m = \frac{2\pi\hbar c n}{eB}.$$

For fixed density n, the levels of B that cause jumps in the conductance are

$$B = \frac{2\pi\hbar c n}{em}. \tag{12.56}$$

This is known as the integer quantum Hall effect. It corresponds to an integral number of levels being exactly filled. From the constraint on magnetic flux given in the discussion of the Aharanov Bohm effect in Eq. (3.33) one can see that these dips in the conductivity happen when the area per electron multiplied by the magnetic field give a magnetic flux corresponding to some integer factor of the relevant quantized units of flux.

The behavior of the conductivity as a function of the magnetic field is remarkable. The conductance is a property of the material described by a tensor, σ_{ij},

$$J_i = \sigma_{ij} E_j. \tag{12.57}$$

This describes the current per unit length in the i direction, as this is a two-dimensional system, in response to a small electric field in the j direction. It should depend on the strength of the magnetic field, which is in the z direction. For free particles in a constant magnetic field $B\hat{z}$, the addition of an electric field $E\hat{x}$ induced a charge to move with an average velocity Ec/B in the y direction as derived in Eq. (3.23). Thus, we expect an off-diagonal component in the conductivity tensor, $\sigma_{xy} \neq 0$, if a magnetic field is present. Additionally, one expects diagonal elements given the presence of scattering in the medium.

It is not too surprising that there are dips in the diagonal components, σ_{xx}, when the applied magnetic field corresponds to an exactly filled shell. This explains the peaks in the longitudinal resistivity, $\rho = \sigma^{-1}$, seen in Fig. 12.2. Indeed, the sharp peaks are located at points where the

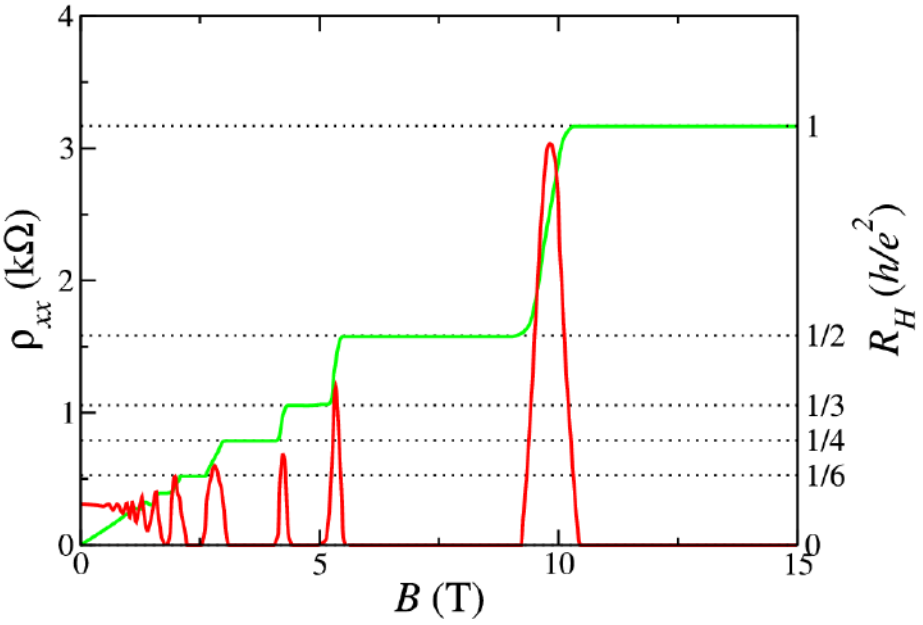


Figure 12.2: Experimental measurements of the diagonal(red) and off-diagonal resistivities (green) in the quantum Hall effect. The off-diagonal resistivity appears at quantized values of $2\pi\hbar/e^2$. Figure from http://oer.physics.manchester.ac.uk/AQM2/Notes/Notes-4.4.html.

magnetic field is adjusted to exactly fill a shell. But it is perhaps surprising to see the longitudinal resistivity vanish between thresholds.

The plateaus in the experimental measurement of the off-diagonal elements in Fig. 12.2 are what are truly surprising. The values of the conductivity at the thresholds can be motivated by simple arguments. If one assumes that the velocity of the electrons correspond to that of freely moving charges,

$$J_y = env_y = en\frac{E_x c}{B}, \tag{12.58}$$

then assigns the magnetic field to correspond to the threshold m defined in Eq. (12.55),

$$\begin{aligned} J_y &= \frac{e^2}{2\pi\hbar}E_x, \\ \sigma_{xy} &= \frac{e^2}{2\pi\hbar}m, \\ \rho_{xy} &= \frac{2\pi\hbar}{me^2}. \end{aligned} \tag{12.59}$$

Although these arguments motivate the off-diagonal resistivity at each threshold of a new Landau level, they fail to describe why the value stays constant between thresholds. The 1985 Nobel Prize was awarded to Klaus von Klitzing for discovery of the quantum Hall effect.

The fractional quantum Hall effect refers to conductance minima for fractional, e.g. one third, values of m. This is much more difficult to explain, but is related to the sharing of electrons

between adjacent orbitals, similar to the sharing of electrons in the covalent bonds in solids. The 1988 Nobel prize was awarded to Robert Laughlin, Horst Störmer and Daniel Tsui for its discovery.

12.16 Exercises

1. Consider the two electron holes in the p-shell of a neutral oxygen atom.
   (a) What is the $\boldsymbol{L-S-J}$ of the ground state.
   (b) If the atom is in a magnetic field of 0.01 Tesla, find the magnetic energies of the originally degenerate $\mathbf{2}\boldsymbol{J}+\mathbf{1}$ states.

2. One electron moves in a one-dimensional system and feels the interaction of two atoms. Approximate the interaction between the electrons and the atoms with the potential
$$\boldsymbol{V(x-R)=-\beta\delta(x-R)},$$
where $\boldsymbol{R}$ is the position of an atom. Use the adiabatic approximation to answer the following questions.
   (a) Given the two atoms are separated by a distance $\boldsymbol{r}$, find a transcendental equation relating $\boldsymbol{k}$ and $\boldsymbol{r}$ where the electronic binding energy is $\hbar^{\mathbf{2}}\boldsymbol{k}^{\mathbf{2}}/(\mathbf{2}\boldsymbol{m})$.
   (b) Find the potential between the two atoms at small $\boldsymbol{r}$,
$$\boldsymbol{V(r\to 0)\sim V(r=0)-\alpha r},$$
   that is, find $\boldsymbol{V(r=0)}$ and $\boldsymbol{\alpha}$. Do this by expanding the transcendental equation in terms of $\boldsymbol{r}$. Hint: First, find $\boldsymbol{V(r=0)}$ by solving the transcendental equation with $\boldsymbol{r=0}$. Take derivatives of the transcendental equation with respect to $\boldsymbol{r}$, then solve for $\boldsymbol{dk/dr}$ at $\boldsymbol{r=0}$, and finally find $\boldsymbol{dE/dr}$ to obtain $\boldsymbol{\alpha}$.
   (c) Find the potential between the two atoms at large $\boldsymbol{r}$,
$$\boldsymbol{V(r\to\infty)=-\gamma\exp(-2k_\infty r)},$$
   that is, find γ. Hint: Use first order perturbation theory, assuming the unperturbed wave function is the bound state of one well, and the perturbation is the interaction with the second well.

3. Consider a particle of mass $\boldsymbol{M}$ and charge $\boldsymbol{e}$ moving in the $\boldsymbol{x-y}$ plane under the influence of a magnetic field in the $\boldsymbol{z}$ direction. Ignore motion in the $\boldsymbol{z}$ direction.
   (a) Show that the vector potential,
$$\vec{\boldsymbol{A}}=\frac{\boldsymbol{B}}{\mathbf{2}}\left(\boldsymbol{x}\hat{\boldsymbol{y}}-\boldsymbol{y}\hat{\boldsymbol{x}}\right),$$
   describes a magnetic field in the $\boldsymbol{z}$ direction.
   (b) Express the vector potential in cylindrical coordinates, that is in terms of $\boldsymbol{r}$, $\boldsymbol{\phi}$, $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{\phi}}$.

(c) Write the Schrödinger equation,

$$\frac{\left(\vec{\boldsymbol{P}} - e\vec{\boldsymbol{A}}/c\right)^2}{2m}\psi(r,\phi) = E\psi(r,\phi),$$

in cylindrical coordinates.

(d) Show that $\boldsymbol{L_z}$ commutes with the Hamiltonian.

(e) Assuming the solution is an eigenstate of $\boldsymbol{L_z}$ with eigenvalue $m\hbar$,

$$\psi(r,\phi) = e^{im\phi}\xi_m(r),$$

rewrite the Schrödinger equation for $\xi_m(r)$.

(f) **Extra Credit**: Solve for $\xi_m(r)$ and the eigenenergies for the case where $m = 0$.

4. Consider a surface with 10 electrons per μm^2. Lowering the magnetic field, at what magnetic field (in Tesla) do you find the first dip in conductivity due to the quantum Hall effect?

13 Relativistic Quantum Mechanics

13.1 The Klein-Gordon Equation

In relativistic kinematics energy and momentum appear on an equal footing, as do position and time. For instance,

$$\boldsymbol{E^2 - p^2c^2 = m^2c^4.} \tag{13.1}$$

Rewriting energy and momentum as operators, one obtains the Klein-Gordon equation,

$$\left(-\hbar^2\frac{\boldsymbol{\partial^2}}{\boldsymbol{\partial t^2}} + \hbar^2\boldsymbol{c^2\nabla^2} - \boldsymbol{m^2c^4}\right)\boldsymbol{\phi(\vec{r},t) = 0.} \tag{13.2}$$

In this version, the wave function ϕ can be a complex number. There are instances where one uses a real field, but here we use a complex field so that we can find solutions that are plane waves. It differs from the Schrödinger equation in that time appears in a second derivative.

The complex Klein-Gordon equation has plane wave solutions,

$$\begin{aligned}\boldsymbol{\phi_{\vec{p}}(\vec{r})} &\boldsymbol{\sim e^{i\vec{p}\cdot\vec{r}/\hbar - iE_\pm(p)t/\hbar},} \\ \boldsymbol{E_\pm(p)} &\boldsymbol{= \pm\sqrt{p^2c^2 + m^2c^4}}\end{aligned} \tag{13.3}$$

If one expands the energy for small $\boldsymbol{p}$,

$$\boldsymbol{E_+(p) \approx mc^2 + p^2/2m.} \tag{13.4}$$

Thus, in the non-relativistic limit, the energy appears identical to the that of the Schrödinger equation aside from the inclusion of the rest mass energy. If mass were conserved this would merely add a constant to the energy along with an unobservable phase, $\boldsymbol{e^{-iMc^2t/\hbar}}$. However, for cases where particles are generated, the phase plays an important role.

Perhaps the most unusual feature in the solutions is the appearance of negative-energy solutions. The negative-energy solutions can only be understood when one writes field operators in second quantization. The field operators (remember the case for photons) will then be written as

$$\begin{aligned}\boldsymbol{\Phi(\vec{r},t)} &\boldsymbol{= \sqrt{\frac{\hbar c}{V}}\sum_p \frac{1}{\sqrt{E_p}}\left(e^{ip\cdot r/\hbar}b^\dagger_{\vec{p}} + e^{-ip\cdot r/\hbar}a_{\vec{p}}\right)} \\ &\boldsymbol{= \sqrt{\frac{\hbar c}{V}}\sum_p \frac{1}{\sqrt{E_p}}e^{-i\vec{p}\cdot\vec{r}/\hbar}\left(e^{iEt/\hbar}b^\dagger_{\vec{p}} + e^{-iEt/\hbar}a_{-\vec{p}}\right).}\end{aligned} \tag{13.5}$$

where $\boldsymbol{p\cdot r \equiv E_pt - \vec{p}\cdot\vec{r}}$. Thus, in the language of field operators, the negative-energy solution is associated with the destruction of antiparticles moving with opposite momentum. We note that if the field were a real field, rather than a complex field, $\boldsymbol{b^\dagger_{\vec{p}}}$ would be replaced by $\boldsymbol{a^\dagger_{\vec{p}}}$ and $\boldsymbol{b_{\vec{p}}}$ by $\boldsymbol{a_{\vec{p}}}$. It only seems natural that a relativistic theory should require the consideration of field operators to make sense, because the creation and destruction of particles is an inherent feature of relativistic treatments.

For non-relativistic field operators we had defined field operators as,

$$\Psi(\vec{r},t) = \frac{1}{\sqrt{V}}\sum_k e^{-iE_k t + i\vec{k}\cdot\vec{r}} a_{\vec{k}}, \tag{13.6}$$
$$\Psi^\dagger(\vec{r},t) = \frac{1}{\sqrt{V}}\sum_k e^{iE_k t - i\vec{k}\cdot\vec{r}} a^\dagger_{\vec{k}}.$$

If the field operators refer to electrons, one can think of $\Psi^\dagger(\vec{r})$ as creating a negative particle at $\vec{r}$, while $\Psi(\vec{r})$ would destroy a negative particle. With relativistic fields, the operator $\Phi^\dagger$ creates a positive particle, but also destroys the negative antiparticle.. This can be considered a particle-antiparticle symmetry that derives from the existence of negative-energy solutions. The same symmetry will result from the Dirac equation, which is the topic of the next chapter.

13.2 Current Conservation

The current density should have the form of a relativistic four-vector. Furthermore, the current density should be conserved,

$$\partial_t j_0(\vec{r},t) + \nabla\cdot\vec{j}(\vec{r},t) = 0, \tag{13.7}$$

or in four-vector notation,

$$\partial_\mu j^\mu(\vec{r},t) = 0. \tag{13.8}$$

Non-relativistically, the charge density j^0 appeared in the form

$$\rho = e\psi^*(\vec{r},t)\psi(\vec{r},t), \tag{13.9}$$

whereas a derivative appeared in the form for the current density

$$\vec{j}(\vec{r},t) = \frac{-ie\hbar}{2m}\left(\psi^*(\vec{r},t)\vec{\partial}\psi(\vec{r},t) - \vec{\partial}\psi^*(\vec{r},t)\psi(\vec{r},t)\right). \tag{13.10}$$

This clearly violates the relativistic spirit where the time and space components of a four vector should appear with similar forms.

As an ansatz, we guess a form for the four-current,

$$j^0(\vec{r},t) = \frac{-ie}{2c}\left(\phi^*(\vec{r},t)\frac{\partial}{\partial t}\phi(\vec{r},t) - \frac{\partial}{\partial t}\phi^*(\vec{r},t)\phi(\vec{r},t)\right) \tag{13.11}$$
$$\vec{j}(\vec{r},t) = \frac{-iec}{2}\left(\phi^*(\vec{r},t)\vec{\partial}\phi(\vec{r},t) - \vec{\partial}\phi^*(\vec{r},t)\phi(\vec{r},t)\right).$$

It is now straight-forward to see that

$$\partial_\mu j^\mu(\vec{r},t) = 0, \tag{13.12}$$

by taking the four divergence of j^μ and then applying the Klein-Gordon equation.

Aside: Covariant-Contravariant Notation: Four vectors, e.g. time and position or energy and momentum, are usually referenced by Greek indices numbering from zero to three. The zero$^{\mathbf{th}}$ component refers to the component which rotate and boost like time, e.g. time or energy, and the other three refer to the components that behave like spatial vectors, e.g. position or momentum. Contravariant vectors are labeled with superscripts, e.g. $\boldsymbol{A^\mu}$, and the values are the actual values of the vectors. Covariant vectors are labeled by subscripts and have their spatial components reversed. The components of $\boldsymbol{p_\mu}$ are $\boldsymbol{E}$, $\boldsymbol{-p_x c}$, $\boldsymbol{-p_y c}$, $\boldsymbol{-p_z c}$. Derivatives behave oppositely,

$$\boldsymbol{\partial_\mu = \frac{\partial}{\partial x^\mu}, \;\; \partial^\mu = \frac{\partial}{\partial x_\mu}.} \tag{13.13}$$

Here, the factors of $\boldsymbol{c}$ are neglected. Once $\boldsymbol{c}$ is included one must worry about whether the position $\boldsymbol{x^\mu}$ refers to $\boldsymbol{(ct, x, y, z)}$ or $\boldsymbol{t, x/c, y/c, z/c}$. The best way to proceed is to ignore factors of $\boldsymbol{c}$ and treat time and position as if they are measured in the same units. This only makes sense, because boosts behave like rotations that mix time and position. Maintaining a different set of units for time and position (or for energy and momentum) makes as much sense as using one set of units for the $\boldsymbol{x}$-component of the position and another for the $\boldsymbol{y}$-component. For the remainder of the chapter factors of $\boldsymbol{c}$ may appear or disappear from expressions. Typically, one ignores factors of $\boldsymbol{c}$ until the final expressions are determined. Then, when actual numbers are needed one makes sure that the various quantities are consistent. For instance, in high-energy nuclear physics distance is often measured in fm, and time is measured in fm/$\boldsymbol{c}$. In this set of units $\boldsymbol{c = 1}$, so all the factors can be ignored. If one ultimately requires knowing time in seconds, one first calculates a numerical value for the time in fm/$\boldsymbol{c}$, then converts to seconds.

The use of Greek indices for four-vectors, and Roman indices for three-vectors, helps alleviate some of the confusion with covariant and contravariant notation. For instance,

$$\boldsymbol{A_{i=2} = A^{\mu=2} = -A_{\mu=2}.} \tag{13.14}$$

Thus, the term $\boldsymbol{A_2}$ is ambiguous, unless one one knows whether one is referring to a four-vector or a three-vector. Dot-products of four vectors,

$$\boldsymbol{A \cdot B = A^\mu B_\mu = A_0 B_0 - A_x B_x - A_y B_y - A_z B_z,} \tag{13.15}$$

are invariant to both boosts and rotations. To distinguish these from the dot products of three-vectors, $\boldsymbol{\vec{A} \cdot \vec{B} = A_x B_x + A_y B_y + A_z B_z}$, the vector (or in other text bold face) are not invoked, i.e. the dot-product of a four-vector is $\boldsymbol{A \cdot B}$, not $\boldsymbol{\vec{A} \cdot \vec{B}}$.

Furthermore, one can see that the look at the form of $\boldsymbol{j^\mu}$ in terms of field operators. One then obtains, after integrating over space,

$$\boldsymbol{\int d^3 r \; j^0(\vec{r}, t) = e \sum_p \left(a^\dagger_{\vec{p}} a_{\vec{p}} - b_{\vec{p}} b^\dagger_{\vec{p}} \right)} \tag{13.16}$$

$$\boldsymbol{\int d^3 r \; \vec{j}(\vec{r}, t) = e c^2 \sum_p \frac{\vec{p}}{E_p} \left(a^\dagger_{\vec{p}} a_{\vec{p}} - b_{\vec{p}} b^\dagger_{\vec{p}} \right)}$$

These look like the usual non-relativistic expressions, but particles and antiparticles contributing with opposite signs.

A peculiar aspect of the expression is that the contribution from antiparticles has the creation and destruction operators reversed, i.e. $bb^\dagger$ rather than the number operator $b^\dagger b$. This implies that an extra charge of -1 is associated with each momentum mode after making the substitution,

$$a^\dagger a - bb^\dagger = a^\dagger a - b^\dagger b - 1. \tag{13.17}$$

This suggests that the vacuum has a charge of -1 associated with each mode, which translates into an infinite negative charge because there are an infinite number of modes.

If one wrote down an expression for the Hamiltonian in terms of field operators (which could be derived from Lagrangians) one would find,

$$\begin{aligned} H &= \sum_p E_p \left(a^\dagger_{\vec{p}} a_{\vec{p}} + b_{\vec{p}} b^\dagger_{\vec{p}} \right) \\ &= \sum_p E_p \left(a^\dagger_{\vec{p}} a_{\vec{p}} + b^\dagger_{\vec{p}} b_{\vec{p}} + 1 \right). \end{aligned} \tag{13.18}$$

Thus, the vacuum has an infinite positve energy associated with it. When we consider the Dirac equation, which is applicable for fermions, the same problem will appear, but with opposite signs. That is, the energy of the vaccum will be $-E_p$ for each mode. This result serves as one of the motivations for super-symmetry, where every bosonic mode has a corresponding fermionic mode.

13.3 Coupling to the Electromagnetic Field

In the Schrödinger equation, incorporating the electromagnetic field could be accomplished by minimal substitution,

$$\begin{aligned} -i\hbar\vec{\partial} &\to -i\hbar\vec{\partial} - e\vec{A}/c \\ i\hbar\frac{\partial}{\partial t} &\to i\hbar\frac{\partial}{\partial t} - eA_0, \end{aligned} \tag{13.19}$$

where Φ is the "scalar" field. However, in relativistic electromagnetism, Φ is the zeroth component of the four vector A^μ.

In relativistic treatments, coupling to the electromagnetic field is also incorporated with the same minimal substitution. One result of coupling to an electromagnetic field is that the charge and current densities are modified,

$$\begin{aligned} j^0(\vec{r},t) &= \frac{ie}{2c}\left(\phi^*(\vec{r},t)\frac{\partial}{\partial t}\phi(\vec{r},t) - \frac{\partial}{\partial t}\phi^*(\vec{r},t)\phi(\vec{r},t) \right) - \frac{e^2 A_0(\vec{r},t)}{\hbar c}\phi^*(\vec{r},t)\phi(\vec{r},t) \\ \vec{j}(\vec{r},t) &= \frac{-iec}{2}\left(\phi^*(\vec{r},t)\vec{\partial}\phi(\vec{r},t) - \vec{\partial}\phi^*(\vec{r},t)\phi(\vec{r},t) \right) - \frac{e^2\vec{A}(\vec{r},t)}{\hbar}\phi^*(\vec{r},t)\phi(\vec{r},t). \end{aligned} \tag{13.20}$$

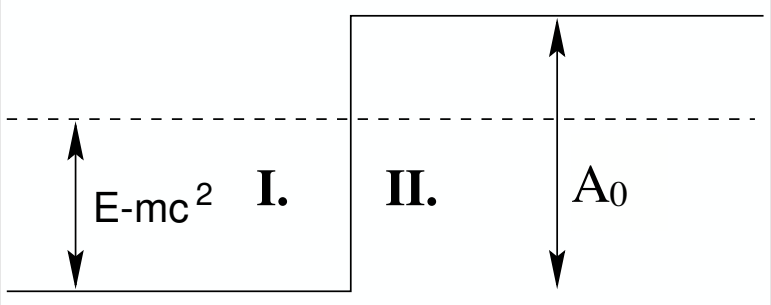
The charge density can change signs if the energy, $\langle i\hbar\partial_t \rangle$, changes sign, or if the electric potential A_0 is altered. Negative charge density can be interpreted as the presence of antiparticles.

The electromagnetic field also affects the energy through minimal substitution,

$$(E - eA_0)^2 = (\vec{P} - e\vec{A})^2 + m^2, \tag{13.21}$$
$$E = eA_0 \pm \sqrt{(\vec{P} - e\vec{A})^2 + m^2}\ .$$

For the case where $\vec{A} = 0$ and I have a simple electromagnetic field, states with a given momentum $\vec{P}$ might change from positive to negative energy depending on the strength and sign of A_0. If a state has negative energy, and if you are searching for the minimum energy state, you might expect the ground state to be populated with infinite numbers of quanta. Interpreting such states can be challenging.

Example 13.1: Reflecting off a Potential Step
A seemingly simple example that becomes surprisingly difficult is the reflection off a potential step, where the potential step is due to jump in A_0. Consider the one-dimensional problem,



$$A_0(x) = \begin{matrix} 0, & x < 0 \\ A_0, & x > 0 \end{matrix}$$

Consider a wave incident from the left, with a reflected and transmitted wave,

$$\phi_I(x) = e^{ikx} - Be^{-ikx}$$
$$\phi_{II}(x) = Ce^{ik'x},$$

where the energy of the particle is

$$E = \sqrt{\hbar^2 k^2 c^2 + m^2 c^4}.$$

Before solving for B and C, one must find k'. Applying the Klein-Gordon equation,

$$(E - eA_0)^2 = \hbar^2 k'^2 c^2 + m^2 c^4,$$
$$k'^2 = \frac{(E - eA_0)^2 - m^2 c^4}{\hbar^2 c^2}.$$

One may then find the ranges for which k'^2 is positive or negative,

1. $k'^2 > 0$ when $0 < eA_0 < E - mc^2$,
2. $k'^2 < 0$ when $E - mc^2 < eA_0 < E + mc^2$,
3. $k'^2 > 0$ when $eA_0 > E + mc^2$.

The first two ranges for A_0 correspond to the usual case for non-relativistic physics. That is, when the barrier is small the particle propagates as a plane wave with a reduced velocity. For larger barriers, the particle is confined to region I and the wave function exponentially dies in region II.
However, the surprise here is that for very large A_0, there again appears to be a solution with the particle moving through the barrier. This surprising result can be understood by calculating the charge and current densities in region II. One finds that the charge density in region II is opposite to that of region I. The solution describes an incoming wave that splits into two outgoing waves, one of particles moving to the left and one of antiparticles moving to the right. This behavior owes itself to the large field. When the voltage difference exceeds twice the mass, it becomes possible to create pairs of particles. Because the antiparticle feels a potential of $-eA_0$, it is perfectly willing to go to into region II, and many do so spontaneously if the step has enough energy to overcome the mass penalty, $2mc^2$. This behavior does not exist should one consider a problem where the regions have different masses rather than different values of eA_0.

13.4 The Dirac Equation

The Dirac equation is another example of a relativistic wave equation. However, the Dirac equation differs in that it describes spin 1/2 particles whereas the Klein-Gordon equation applies to spinless particles.

Dirac was motivated by finding a linear equation that was consistent with the relativistic constraint $H^2 = P^2c^2 + m^2c^4$. By linear, Dirac was looking for an equation that was linear in the derivatives. In order to accomplish this Dirac suggested a matrix equation,

$$\begin{aligned} H &= \alpha_x c p_x + \alpha_y c p_y + \alpha_z c p_z + \beta mc^2 \\ &= c\vec{\alpha}\cdot\vec{p} + \beta mc^2, \end{aligned} \tag{13.22}$$

where α_i and β are Hermitian matrices. Given that one needs to satisfy our knowledge of relativity,

$$H^2 = p^2c^2 + m^2c^4, \tag{13.23}$$

there are constraints on the matrices,

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij}, \beta^2 = 1, \{\alpha_i, \beta\} = 0. \tag{13.24}$$

Two-by-two matrices are insufficient, as one can only comprise three anti-commuting matrices (the Pauli matrices), but four are needed. Three-by-three matrices are also insufficient, but four-by-four matrices are sufficient. The following four-by-four matrices satisfy the conditions,

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}, \ \beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{13.25}$$

Actually, there exist an infinite number of choices as one can transform the matrices by a unitary transformation,

$$\boldsymbol{U\vec{\alpha}U^\dagger}, \tag{13.26}$$

and find a new choice of matrices. The choice above is known as the Dirac representation, which is convenient for massive particles. For highly relativistic particles, it is sometimes convenient to employ the chiral representation,

$$\vec{\boldsymbol{\alpha}} = \begin{pmatrix} \vec{\boldsymbol{\sigma}} & \mathbf{0} \\ \mathbf{0} & -\vec{\boldsymbol{\sigma}} \end{pmatrix}, \ \boldsymbol{\beta} = -\begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \end{pmatrix}. \tag{13.27}$$

We will mainly employ the Dirac representation.

It should be stressed that the four components of the wave function are not a relativistic four vector. They correspond to spin-up, spin-down, and the two corresponding anti-particle solutions. Like any wave function, the components of the wave functions mean nothing in isolation. Only expectations have a physical meaning, and those involve overlaps of vectors. By changing the arbitrarily chosen representation, the physical meaning of the solutions is unchanged.

13.5 Conserved Quantities

The particle current,

$$\boldsymbol{j_0}(\vec{\boldsymbol{r}},\boldsymbol{t}) \equiv \boldsymbol{\psi^\dagger}(\vec{\boldsymbol{r}},\boldsymbol{t})\boldsymbol{\psi}(\vec{\boldsymbol{r}},\boldsymbol{t}), \ \ \vec{\boldsymbol{j}}(\vec{\boldsymbol{r}},\boldsymbol{t}) \equiv \boldsymbol{c\psi^\dagger}(\vec{\boldsymbol{r}},\boldsymbol{t})\vec{\boldsymbol{\alpha}}\boldsymbol{\psi}(\vec{\boldsymbol{r}},\boldsymbol{t}), \tag{13.28}$$

is conserved. One can demonstrate the conservation of the current by taking the four divergence, $\boldsymbol{\partial_\mu j^\mu}$, and by applying the Dirac equation,

$$\begin{aligned} \boldsymbol{H\psi} &= \boldsymbol{i\hbar\frac{\partial}{\partial t}\psi}, \\ \boldsymbol{i\hbar\frac{\partial}{\partial t}\psi + i\hbar c\vec{\alpha}\cdot\vec{\nabla}\psi} &= \boldsymbol{mc^2\beta\psi}, \\ \boldsymbol{-i\hbar\frac{\partial}{\partial t}\psi^\dagger - i\hbar c\vec{\alpha}\cdot\vec{\nabla}\psi^\dagger} &= \boldsymbol{mc^2\beta\psi^\dagger}, \end{aligned} \tag{13.29}$$

to see that

$$\begin{aligned} \boldsymbol{\frac{\partial}{\partial t}j_0 + \nabla\cdot\mathbf{j}} &= \boldsymbol{[(\partial_t + c\nabla\cdot\vec{\alpha})\psi(\vec{r},t)]^\dagger\psi(\vec{r},t) + \psi^\dagger(\vec{r},t)(\partial_t + c\nabla\cdot\vec{\alpha})\psi(\vec{r},t)} \\ &= \boldsymbol{\frac{imc^2}{\hbar}\left[-(\beta\psi(\vec{r},t)^\dagger\psi(\vec{r},t) + \psi^\dagger(\vec{r},t)\beta\psi(\vec{r},t)\right]} \\ &= \mathbf{0}. \end{aligned} \tag{13.30}$$

The properties that $\vec{\boldsymbol{\alpha}}$ and $\boldsymbol{\beta}$ are Hermitian were used.

This contrasts with the Klein-Gordon equation in that no derivatives are used to construct the current-density or the charge-density operators. In the Klein-Gordon equation both quantities had derivatives, and in the Schrödinger equation the current density involves derivatives, but not the charge density.

Now, we consider conservation of angular momentum. Unlike the Schrödinger equation, the Dirac Hamiltonian does not commute with the orbital angular momentum.

$$\begin{aligned}[\boldsymbol{H}, \vec{\boldsymbol{r}} \times \vec{\boldsymbol{p}}]_i &= \boldsymbol{c}[\sum_{\ell} \boldsymbol{\alpha_\ell p_\ell}, \sum_{jk} \boldsymbol{\epsilon_{ijk} r_j p_k}] + \boldsymbol{mc^2}[\boldsymbol{\beta}, \sum_{jk} \boldsymbol{\epsilon_{ijk} r_j p_k}] \\ &= -\boldsymbol{i\hbar c} \sum_{jk} \boldsymbol{\epsilon_{ijk} \alpha_j p_k}, \\ [\boldsymbol{H}, \vec{\boldsymbol{r}} \times \vec{\boldsymbol{p}}] &= -\boldsymbol{i\hbar c} \vec{\boldsymbol{\alpha}} \times \vec{\boldsymbol{p}}.\end{aligned} \tag{13.31}$$

One can also define a spin operator, $\boldsymbol{S} = \hbar\vec{\boldsymbol{\Sigma}}/\boldsymbol{2}$, where

$$\begin{aligned}\boldsymbol{\Sigma_i} &= -\frac{\boldsymbol{i}}{\boldsymbol{2}} \boldsymbol{\epsilon_{ijk} \alpha_i \alpha_j} \\ \vec{\boldsymbol{\Sigma}} &= \begin{pmatrix} \vec{\boldsymbol{\sigma}} & \boldsymbol{0} \\ \boldsymbol{0} & \vec{\boldsymbol{\sigma}} \end{pmatrix} \textbf{ in the Dirac representation.}\end{aligned} \tag{13.32}$$

The spin is also not conserved.

$$\begin{aligned}[\boldsymbol{H}, \boldsymbol{S_i}] &= \sum_j \frac{\hbar \boldsymbol{c p_j}}{\boldsymbol{2}} [\boldsymbol{\alpha_j}, \boldsymbol{\Sigma_i}] \\ &= \sum_{jkl} \frac{-\boldsymbol{i\hbar c p_j}}{\boldsymbol{4}} \boldsymbol{\epsilon_{ikl}} [\boldsymbol{\alpha_i}, \boldsymbol{\alpha_k \alpha_l}] \\ &= \sum_{jk} \frac{-\boldsymbol{i\hbar c p_j}}{\boldsymbol{2}} \left(\boldsymbol{\epsilon_{ijk} \alpha_k} - \boldsymbol{\epsilon_{ikj} \alpha_k}\right) \\ &= \boldsymbol{i\hbar c} \sum_{jk} \boldsymbol{\epsilon_{ijk} \alpha_j p_k}, \\ [\boldsymbol{H}, \vec{\boldsymbol{S}}] &= \boldsymbol{i\hbar c} \vec{\boldsymbol{\alpha}} \times \vec{\boldsymbol{p}}.\end{aligned} \tag{13.33}$$

However, the combination $\vec{\boldsymbol{J}} = \vec{\boldsymbol{r}} \times \vec{\boldsymbol{p}} + \vec{\boldsymbol{S}}$ is conserved,

$$[\boldsymbol{H}, \vec{\boldsymbol{J}}] = \boldsymbol{0}. \tag{13.34}$$

13.6 Solutions for Free Particles

In analogy to the Klein-Gordon equation, the Dirac equation will have both positive-energy and negative-energy solutions. Here, we consider those solution which behave as $\boldsymbol{e^{i\vec{p}\cdot r}}$. The positive-energy solutions with a given momentum, $\vec{\boldsymbol{p}}$, are referred to as $\boldsymbol{u_s(\vec{p})}$ while the negative-energy solutions are referred to as $\boldsymbol{v_s(-\vec{p})}$. The label $\boldsymbol{s}$ refers to the spin.

$$\begin{aligned}\boldsymbol{E_p u(\vec{p})} &= \vec{\boldsymbol{\alpha}} \cdot \vec{\boldsymbol{p}} \boldsymbol{c u(\vec{p})} + \boldsymbol{\beta mc^2 u(\vec{p})} \\ -\boldsymbol{E_p v(-\vec{p})} &= \vec{\boldsymbol{\alpha}} \cdot \vec{\boldsymbol{p}} \boldsymbol{c v(-\vec{p})} + \boldsymbol{\beta mc^2 v(-\vec{p})}.\end{aligned} \tag{13.35}$$

The momentum labels on the negative-energy solutions were labeled with the opposite momenta because once one switches to the language of fieled operators they would correspond to the destruction of anti-particles with momentum $-\vec{p}$ rather than the creation of particles of momentum $\vec{p}$.

To find the solutions, first find the solutions for $p = 0$. The solutions are then

$$u_\uparrow\ (p=0) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad u_\downarrow(p=0) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \tag{13.36}$$

$$v_\downarrow(p=0) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_\uparrow\ (p=0) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Note that the solutions v are labeled with a spin index opposite to the eigenvalue of Σ_z. Again, this is because the solution will correspond to the destruction of an antiparticle. It is easy to check that these solutions are eigenstates of the Dirac equation with eigenvalues $\pm mc^2$.

Finding solutions for non-zero momentum can be accomplished by multiplying the zero-momentum solutions by the operator $\pm E_p + \vec{\alpha}\cdot\vec{p}c + mc^2\beta$. This results in a solution to the Dirac equation because

$$(E_p - \vec{\alpha}\cdot\vec{p}c - \beta mc^2)(E_p + \vec{\alpha}\cdot\vec{p}c + \beta mc^2) = 0. \tag{13.37}$$

We define the solution $u_s(\vec{p})$ to be

$$u_s(\vec{p}) = \frac{E_p + \vec{\alpha}\cdot\vec{p}c + \beta mc^2}{\sqrt{2mc^2(mc^2+E_p)}} u_s(p=0), \tag{13.38}$$

which will also be a solution to the Dirac equation.

$$(E - \vec{\alpha}\cdot\vec{p}c - \beta mc^2)u(\vec{p}) = 0. \tag{13.39}$$

The square root in the denominator was chosen to result in the normalization

$$u_s^\dagger(p)u_s(p) = \frac{E_p}{mc^2}. \tag{13.40}$$

The choice of normalization is motivated by the fact that $u^\dagger u$ turns out to be related to the zeroth component of the current, which must transform like the zeroth part of a four vector $(E, \vec{p}c)/m$.

As an example, we find $u_\uparrow\ (\vec{p})$ where $\vec{p}$ is along the z axis. For instance,

$$u_\uparrow\ (p_z) = \frac{1}{\sqrt{2mc^2(E_p+mc^2)}} \begin{pmatrix} E_p+mc^2 & 0 & p_zc & 0 \\ 0 & E_p+mc^2 & 0 & -p_zc \\ p_zc & 0 & E_p-mc^2 & 0 \\ 0 & -p_zc & 0 & E_p-mc^2\ . \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{13.41}$$

$$= \frac{1}{\sqrt{2mc^2(E_p+mc^2)}} \begin{pmatrix} E_p+mc^2 \\ 0 \\ p_zc \\ 0 \end{pmatrix}.$$

At non-zero momentum the positive-energy solutions have a mixture of upper (top two) and lower (bottom two) components, with the upper components dwarfing the lower components when the motion is non-relativistic.

Example 13.2: The Infinite Square Well with a Scalar Potential
One can consider a binding potential as having a position dependent mass, i.e. a scalar field $\Phi(x)$. Consider a spin-1/2 particle moving in one dimension with zero mass inside the well, but infinite mass outside the well. The Dirac equation and well are

$$\begin{aligned} E\psi(x) &= i\hbar c\alpha_x \partial_x \psi(x) + \Phi(x)\beta\psi(x), \\ \Phi(x) &= \begin{cases} Mc^2, & x < 0 \\ 0, & 0 < x < L \\ Mc^2, & x > L \end{cases} \\ M &\to \infty. \end{aligned}$$

1. For $x < 0$ and $x > L$ show that the solution behaves as $e^{-Mc|x|/\hbar}$ and $e^{-Mc|x-L|/\hbar}$ respectively.
2. Show that the boundary condition at the surface is

$$i\beta\vec{\alpha}\cdot\hat{n}\psi = \psi.$$

   Note this means that the wave functions are not vanishing at the boundary.
3. Using these boundary conditions find the ground state energy.

Solution:
(1.) Outside the well one can ignore the term in the Dirac equation proportional to E as that is finite, and one must cancel the infinite terms. Ignoring the finite terms the equation becomes

$$\begin{aligned} (-i\hbar c\alpha_x\partial_x - Mc^2\beta)^2\psi(x) &= 0, \\ -c^2\hbar^2\partial_x^2\psi(x) + M^2c^4\psi(x) &= 0. \end{aligned}$$

The term with ∂_x can be arbitrarily large if the wave function changes a finite amount over an infinitesimal distance, so it cannot be discarded. Here, the facts that $\alpha_x^2 = \beta^2 = 1$ and $\{\alpha_x, \beta\} = 0$ were invoked. The solution must then behave as $e^{-Mc|x|/\hbar}$.
(2.) One can integrate the Dirac equation above,

$$\begin{aligned} \int_{-\infty}^{0} dx[-i\hbar c\alpha_x\partial_x\psi(x)] &= -Mc^2\beta\int_{-\infty}^{0} dx\,\psi(x) \\ -i\hbar c\alpha_x\psi(x=0) &= -Mc^2\beta\int_{-\infty}^{0} dx\,\psi(x=0)e^{-Mc|x|/\hbar}, \\ i\alpha_x\psi(x=0) &= \beta\psi(x=0). \end{aligned}$$

Or more generally, $i\beta\vec{\alpha}\cdot\hat{n}\psi = \psi$ where $\hat{n}$ is the unit surface vector point outside the well. Thus, on the l.h.s. of the well $i\beta\alpha_x\psi(x=0) = \psi(x=0)$ and on the r.h.s. $i\beta\alpha_x\psi(x=L) = -\psi(x=L)$.

(3.) Inside the well the particles are massless and it is easiest to consider the problem in the chiral representation. In that case

$$\boldsymbol{\alpha_x} = \begin{pmatrix} \boldsymbol{\sigma_x} & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{\sigma_x} \end{pmatrix}$$

$$\boldsymbol{\psi_{+,k}(x)} = \frac{\mathbf{1}}{\sqrt{\mathbf{2}}} \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} \boldsymbol{e^{ikx}}, \quad \boldsymbol{\psi_{+,-k}(x)} = \frac{\mathbf{1}}{\sqrt{\mathbf{2}}} \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{1} \\ \mathbf{1} \end{pmatrix} \boldsymbol{e^{-ikx}}.$$

These are two positive-energy solutions for particles with spins in the $\boldsymbol{x}$ direction moving to the left and right respectively. These solutions are also eigenstates of the spin operator $\boldsymbol{\Sigma_x}$, with eigenvalue $\boldsymbol{+1}$. There are also another two solutions that are eigenstates of $\boldsymbol{\Sigma_x}$ with eigenvalues $\boldsymbol{-1}$. Because $\boldsymbol{\Sigma_x}$ commutes with the one-dimensional Hamiltonian (both inside and outside the well), we can choose one set of solutions.
To solve the B.C. one consider a linear combination,

$$\boldsymbol{\psi(x) = A\psi_{+,k}(x) + B\psi_{+,-k}(x).}$$

To consider the B.C. one can write the matrix

$$\boldsymbol{-i\beta\alpha_x} = \boldsymbol{i} \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma_x} \\ \boldsymbol{-\sigma_x} & \mathbf{0} \end{pmatrix},$$

using the definitions of the Dirac matrices from the chiral representation in Eq. (13.27). At $\boldsymbol{x = 0}$, the B.C. can be applied,

$$\begin{aligned} \boldsymbol{i\beta\alpha_x \{A\psi_{+,k}(x=0) + B\psi_{+,-k}(x=0)\}} &= \boldsymbol{A\psi_{+,k}(x=0) + B\psi_{+,-k}(x=0),} \\ \boldsymbol{iA\psi_{+,-k}(x=0) - iB\psi_{+,k}(x=0)} &= \boldsymbol{A\psi_{+,k}(x=0) + B\psi_{+,-k}(x=0).} \end{aligned}$$

This will be a solution of the B.C. if

$$\boldsymbol{B/A = i.}$$

Repeating the considerations at $\boldsymbol{x = L}$ gives

$$\boldsymbol{B/A = -ie^{2ikL}.}$$

The negative sign appeared due to the fact that the unit surface vector switched directions at the other end of the well. The conditions on $\boldsymbol{k}$ are thus

$$\begin{aligned} \boldsymbol{e^{2ikL}} &= \boldsymbol{-1} \\ \boldsymbol{kL} &= \boldsymbol{(n + 1/2)\pi}, \quad \boldsymbol{n = 0, 1, 2 \cdots} \end{aligned}$$

This differs from the Schrödinger equation result, $\boldsymbol{kL = n\pi}$. The ground-state momentum is half the value of the Schrödinger case, but the separation of the values of $\boldsymbol{k}$ are the same. One consequence of this is that if one considers a macroscopic Fermi gas of such particles, there is no penalty is splitting the well into two wells, and thus there is no surface energy.

13.7 Non-Relativistic Limit and the g Factor

Interaction with the electromagnetic field can be incorporated by replacing $\vec{p}$ with $\vec{p} - e\vec{A}/c$ and $i\hbar\partial/\partial t$ with $i\hbar\partial/\partial t - eA_0$, where A_0 is the electric potential.

Using the Dirac representation, one can write the four components in terms of two two-component vectors ϕ and χ, where ϕ and χ refer to the upper/lower components.

$$\psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix}. \tag{13.42}$$

One can now write the Dirac equation as two equations for ϕ and χ.

$$i\hbar\frac{\partial\phi}{\partial t} = c\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}\right)\cdot\vec{\sigma}\chi + (A_0 + mc^2)\phi, \tag{13.43}$$

$$i\hbar\frac{\partial\chi}{\partial t} = c\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}\right)\cdot\vec{\sigma}\phi + (eA_0 - mc^2)\chi. \tag{13.44}$$

In the non-relativistic limit the lower components will be small. Given that the energy is dominated by the mass, can make the approximation

$$i\hbar\frac{\partial}{\partial t}\chi \approx mc^2\chi, \tag{13.45}$$

which results in a simple substitution for χ using Eq. (13.44).

$$2mc^2\chi \approx c\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}\right)\cdot\vec{\sigma}\phi. \tag{13.46}$$

Substituting this expression into Eq. (13.43),

$$i\hbar\frac{\partial}{\partial t}\phi = \frac{1}{2m}\left[\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}\right)\cdot\vec{\sigma}\right]^2\phi + (eA_0 + mc^2)\phi. \tag{13.47}$$

If the gradient and vector potential operators commuted with one another, one would use the anti-commutation relations to obtain the usual kinetic energy piece in the Schrödinger equation from squaring the $\vec{p} - e\vec{A}/c$ term and employing the anti-commutation relation. However, taking account of such terms yields an extra term.

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\phi &= \frac{1}{2m}\left[\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}\right)\right]^2\phi + (eA_0 + mc^2)\phi + \frac{e\hbar}{2mci}\sum_{i\neq j}\sigma_i\sigma_j(\partial_i A_j + A_i\partial_j)\phi \\ &= \frac{1}{2m}\left[\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}\right)\right]^2\phi + (eA_0 + mc^2)\phi + \frac{e\hbar}{2mci}\sum_{i\neq j}\sigma_i\sigma_j(\partial_i A_j - A_j\partial_i)\phi \\ &= \frac{1}{2m}\left[\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}\right)\right]^2\phi + (eA_0 + mc^2)\phi - \frac{e\hbar}{2mc}\vec{\Sigma}\cdot\vec{B}\phi. \end{aligned} \tag{13.48}$$

Because $\hbar\vec{\Sigma} = 2\vec{S}$, the last term explains why the g factor of the electron is 2. Previously, this had been inserted by hand. However, the Dirac equation requires that the g factor be exactly two (though there are small perturbative corrections). Note that the g factor of the proton and neutron are not equal to two, because they are composite particles. Being composite particles, the spin of the proton is composed of contributions of internal orbital angular momenta of the quarks and gluons.

The only difference between the expression above and the usual Schrödinger equation is in the additional term mc^2. However, this merely adds a constant to the energy as long as mass is conserved, and does not affect any observable in the non-relativistic limit. This rapid oscillation was referred to as "zitterbewegung" (trembling motion) by Schrödinger, https://en.wikipedia.org/wiki/Zitterbewegung.

13.8 The Spin-Orbit Interaction

One can make a rigorous expansion of Eq.s (13.43) and (13.44) by iterating Eq. (13.44),

$$\begin{aligned}\chi &= \frac{1}{2mc}\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\mathbf{A}\right)\cdot\vec{\sigma}\phi - \frac{1}{2mc^2}\left(i\hbar\frac{\partial}{\partial t} - mc^2 - eA_0\right)\chi \\ &= \frac{1}{2mc}\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\mathbf{A}\right)\cdot\vec{\sigma}\phi - \frac{1}{2mc^2}\left(i\hbar\frac{\partial}{\partial t} - mc^2 - eA_0\right) \\ &\quad \left\{\frac{1}{2mc}\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\mathbf{A}\right)\cdot\vec{\sigma}\phi \quad - \frac{1}{2mc^2}\left(i\hbar\frac{\partial}{\partial t} - mc^2 - eA_0\right)\chi\right\}.\end{aligned} \tag{13.49}$$

One can repeat the iterative substitution and find an expression for χ to arbitrary power in $1/m$. Once a satisfactory level for χ has been found, it may be substituted into Eq. (13.43) to obtain a wave equation for ϕ.

If one pursues the expansion one step beyond what was done previously with the magnetic field to find the g factor of the electron, one finds (with a substantial amount of algebra) that the approximate Hamiltonian for the upper components ϕ has the extra terms,

$$\begin{aligned}\delta H &= -\frac{p^4}{8m^3c^2} + i\frac{e\hbar}{8m^2c^2}[(\vec{\sigma}\cdot\vec{E})(\vec{\sigma}\cdot\vec{p}) - (\vec{\sigma}\cdot\vec{p})(\vec{\sigma}\cdot\vec{E})] + \frac{e\hbar^2}{8m^2c^2}\nabla^2 A_0, \\ &= -\frac{p^4}{8m^3c^2} - \frac{e\hbar}{4m^2c^2}[\vec{\sigma}\cdot(\vec{E}\times\vec{p}]) + \frac{e\hbar^2}{8m^2c^2}\nabla^2 A_0,\end{aligned} \tag{13.50}$$

where $\vec{E}$ is the electric field. Here, we have kept terms that are of order $1/m^2$ of the usual non-relativistic terms, noting that $p^2/2m$ is one of the non-relativistic terms. The first term above is merely the next-order expansion of $E = \sqrt{p^2c^2 + m^2c^4}$, while the second term is the spin-orbit term. To show that the second term is the spin-orbit term, we note that

$$\vec{E}\times\vec{p} = \frac{E(r)}{r}\vec{r}\times\vec{p}, \tag{13.51}$$

for a radial electric field $\vec{E}$. The final term in Eq. (13.50) is known as the Darwin term (derived by Charles Galton Darwin, the grandson of the more famous Charles Darwin). For the Coulomb

potential, $A_0 = -Ze/r$, it can be re-expressed as

$$H_{\text{Darwin}} = \frac{Z\hbar^2 e^2}{8m^2c^2}\delta^3(\vec{r}), \tag{13.52}$$

and thus appears mainly for $s-$ wave terms.

Finally, it is noted that one may go through the same exercise without any electromagnetic field, but instead with a position dependent mass $m(r)$. One then finds that the spin-orbit term looks identical except that

$$e\vec{E} \rightarrow \frac{\partial m(r)c^2}{\partial r}. \tag{13.53}$$

This has the opposite contribution for an attractive interaction. This is important for understanding nuclear physics where the spin-orbit interaction is surprisingly large. It can be explained by an attractive scalar interaction (like a position dependent mass) and a repulsive vector interaction (similar to a Hydrogen atom, but with opposite sign because it is repulsive). The interactions cancel each other out to a large degree as far as the binding energies are concerned but the contributions from the spin-orbit terms add together.

The Dirac equation is one of the great triumphs of twentieth century physics. Motivated by aesthetic considerations, several previously ad-hoc assumptions fall out naturally: the g factor, particle-antiparticle symmetry, the spin-orbit coupling, etc. When combined with the relativistic coupling to the quantum electromagnetic field, incredibly accurate calculations can be made of $g-2$ using perturbation theory. But, this is material for another course.

13.9 More Notation: γ Matrices

To more clearly demonstrate the covariant nature of the equations, it is common to define the Dirac γ matrices.

$$\gamma^0 \equiv \beta, \;\; \vec{\gamma} \equiv \beta\vec{\alpha}. \tag{13.54}$$

The four matrices γ^μ then transform into one another as a four-vector. The three space-like γ matrices are anti-Hermitian, while γ^0 is Hermitian. The convenient covariant behavior of the γ matrices comes from the property,

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}, \tag{13.55}$$

where $g^{\mu\nu}$ is the metric tensor

$$g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{13.56}$$

This form confirms that γ^μ was a four-vector.

A very common notation is

$$\not{p} \equiv p^\mu \gamma_\mu, \tag{13.57}$$

which means that the Dirac equation (multiplied by γ_0) can be expressed as

$$(\not{p} + m)\psi = 0. \tag{13.58}$$

This has the attraction that $\not{p}$ is a scalar.

One can write the current as

$$j^\mu = \bar{\psi}\gamma^\mu\psi, \tag{13.59}$$

where

$$\bar{\psi} \equiv \psi^\dagger \gamma_0. \tag{13.60}$$

This seems a bit peculiar, but $\bar{\psi}\psi$ is a scalar rather than $\psi^\dagger\psi$ which transforms like the zeroth component of the four current. To see that $\bar{\psi}\psi$ is a scalar under boosts, we recall Eq. (13.38) where the operator used to boost the solution from $\vec{p} = 0$ to $\vec{p}$ was

$$B = \frac{E + \vec{\alpha}\cdot\vec{p} + \beta m}{\sqrt{2m(m+E_p)}}, \tag{13.61}$$

which is Hermitian. Next, we consider the boosted value of $\bar{\psi}\psi$ where the boosted wave function is ψ and the $p = 0$ solution is ψ_0,

$$\begin{aligned}
\bar{\psi}\psi &= \psi_0^\dagger B^\dagger \beta B \psi_0 \\
&= \psi_0^\dagger \frac{E + \vec{\alpha}\cdot\vec{p} + \beta m}{\sqrt{2m(m+E_p)}} \beta \frac{E\vec{\alpha}\cdot\vec{p} + \beta m}{\sqrt{2m(m+E_p)}} \psi_0 \\
&= \frac{1}{2m(m+E_p)} \psi_0^\dagger \beta (E - \vec{\alpha}\cdot\vec{p} + \beta m)(E\vec{\alpha}\cdot\vec{p} + \beta m)\psi_0 \\
&= \frac{1}{2m(m+E_p)} \psi_0^\dagger \beta (E^2 - p^2 + m^2 + 2mE\beta - 2m\beta\vec{\alpha}\cdot\vec{p})\psi_0.
\end{aligned} \tag{13.62}$$

For the at-rest solution $\psi_0^\dagger\beta\alpha_i\psi_0 = 0$ because α mixes upper and lower components, and $\bar{\psi}_0\psi = \psi_0^\dagger\psi$. Substituting $E^2 = p^2 + m^2$ then gives

$$\bar{\psi}\psi = \bar{\psi}_0\psi_0. \tag{13.63}$$

Though this is not a proof, it certainly shows that $\bar{\psi}\psi$ a Lorentz invariant. Similarly from our study of the currents, $\bar{\psi}\gamma^\mu\psi$ behaves as a four vector.

To more rigorously see the Lorentz structure of all the terms and more consistently derive the expressions for the currents, it is best to start with the Lagrangian density,

$$\mathcal{L} = \bar{\psi}(\not{p} + m)\psi, \tag{13.64}$$

then use Noether's theorem to derive the equations of motion (Dirac's equation) and conserved currents. The Lagrangian density above is manifestly boost-invariant, as it should be. To build a scalar, one need only require that $\bar{\psi}$ and ψ surround operators constructed from four-vectors where all the Lorentz indices are contracted. The four matrices γ^μ indeed behave like four vectors under boosts.

13.10 Dirac Field Operators

For the non-relativistic case field operators were defined as

$$\Psi_i(\vec{r},t) = \frac{1}{\sqrt{V}}\sum_{\vec{p}}\sum_s u_{s,i}(\vec{p})e^{-iE_pt/\hbar+i\vec{p}\cdot\vec{r}/\hbar}b_{s,\vec{p}}. \tag{13.65}$$

Here, the two-component vector $u_i(s)$ describes how the spin s is projected onto the chosen basis. For example $i=1$ might refer to spin-up along the z axis and $i=2$ to spin-down. The spin label s might refer to some other basis, such as along the direction of $\vec{p}$, or perhaps along another axis. The field operator $\Psi_i(\vec{r})$ lowers the charge at position $\vec{r}$ by destroying a particle of spin projection i, whereas $\Psi^\dagger(\vec{r})$ creates a particle at $\vec{r}$. The operators $b_{s,\vec{p}}$ destroy particles of momentum $\vec{p}$ with spin s. Because the basis for the spins when referencing momentum states, labeled s, might differ from the basis used for coordinate space, labeled i, the two-by-two matrix $u_{s,i}(\vec{p})$ is simply the representation of a basis transformation which might differ for each momentum value $\vec{p}$.

For the Dirac case, the field operator will have four components, with $i=1\cdots4$. The extra components account for the existence of anti-particles.

$$\Psi_i(\vec{r},t) = \frac{1}{\sqrt{V}}\sum_{\vec{p}}\sqrt{\frac{m}{E_p}}\sum_s\left(u_{s,i}(\vec{p})e^{-iE_pt/\hbar+i\vec{p}\cdot\vec{r}/\hbar}b_{s,\vec{p}} + v_{s,i}(\vec{p})e^{iE_pt/\hbar-i\vec{p}\cdot\vec{r}/\hbar}d^\dagger_{s,\vec{p}}\right). \tag{13.66}$$

The vectors $u_{s,i}(\vec{p})$ are normalized as

$$\begin{aligned} u_s^\dagger(\vec{p})u_{s'}(\vec{p}) = v_s^\dagger(\vec{p})v_{s'}(\vec{p}) &= (E_p/m)\delta_{ss'}, \\ u_s^\dagger(\vec{p})v_s(-\vec{p}) = v_s^\dagger(\vec{p})u_s(-\vec{p}) &= 0. \end{aligned} \tag{13.67}$$

$\sum_i u^\dagger_{s,i}(\vec{p})u_{s',i}(\vec{p}) = (E_p/m)\delta_{ss'}$. Here the operator $d_s^\dagger(\vec{p})$ creates an anti-particle, which is the same as destroying the negative-energy state of momentum $-\vec{p}$. Note that the overlaps of u_s with v_s are evaluated at opposite momenta. This is to ensure that any two operators that are solutions of a given momentum $\vec{p}$, i.e. the coordinate-space dependence is $e^{+i\vec{p}\cdot\vec{r}}$ are orthonormal.

The charge density operator $\Psi_i^\dagger(\vec{r},t)\Psi_i(\vec{r},t)$ can be integrated to yield the total charge,

$$\begin{aligned} \int d^3r\ \Psi_i^\dagger(\vec{r},t)\Psi_i(\vec{r},t) &= \int d^3r \sum_{\vec{p},s,\vec{p}',s'} \frac{m}{E_pV} \\ &\Big\{u^\dagger_{s,i}(\vec{p})u_{s',i}(\vec{p}')e^{i(\vec{p}-\vec{p}')\cdot\vec{r}/\hbar}b^\dagger_{s,\vec{p}}b_{s',\vec{p}'} + v^\dagger_{s,i}(\vec{p})v_{s',i}(\vec{p}')e^{i(\vec{p}'-\vec{p})\cdot\vec{r}/\hbar}d_{s,\vec{p}}d^\dagger_{s',\vec{p}'} \\ &+u^\dagger_{s,i}(\vec{p})v_{s',i}(\vec{p}')e^{i(\vec{p}+\vec{p}')\cdot\vec{r}/\hbar}b^\dagger_{s,\vec{p}}d^\dagger_{s',\vec{p}'} + v^\dagger_{s,i}(\vec{p})u_{s',i}(\vec{p}')e^{-i(\vec{p}+\vec{p}')\cdot\vec{r}/\hbar}d_{s,\vec{p}}b_{s',\vec{p}'}\Big\} \\ &= \sum_{\vec{p},s}\left\{b^\dagger_{s,\vec{p}}b_{s,\vec{p}} + d_{s,\vec{p}}d^\dagger_{s,\vec{p}}\right\} \\ &= \sum_{\vec{p},s}\left\{b^\dagger_{s,\vec{p}}b_{s,\vec{p}} - d^\dagger_{s,\vec{p}}d_{s,\vec{p}} + 1\right\}. \end{aligned} \tag{13.68}$$

The vacuum thus has a contribution to the charge density from each spin and for each value of the momentum.

13.11 The Dirac Equation and the Mass Term

The Hamiltonian for the Dirac equation for massless particles can be written in terms of creation and destruction operators,

$$\boldsymbol{H_0} = \sum_{\vec{k}} \hbar \boldsymbol{k c} (\boldsymbol{b^\dagger_{\vec{k}} b_{\vec{k}}} + \boldsymbol{d^\dagger_{-\vec{k}} d_{-\vec{k}}}), \tag{13.69}$$

where $\boldsymbol{b^\dagger}$ and $\boldsymbol{d^\dagger}$ correspond to creation and destruction operators for particles and antiparticles respectively. For this discussion we ignore the sums over spin or the vacuum energy.

If the mass term is treated like a perturbation, it can also be written in terms of the same creation/destruction operators,

$$\begin{aligned} \boldsymbol{V} &= \boldsymbol{mc^2} \int \boldsymbol{d^3 r}\ \bar{\Psi}(\vec{r}) \Psi(\vec{r}) \\ &= \sum_{\vec{k}} \boldsymbol{mc^2} (\boldsymbol{b^\dagger_{\vec{k}} d^\dagger_{-\vec{k}}} + \boldsymbol{d_{-\vec{k}} b_{\vec{k}}}). \end{aligned} \tag{13.70}$$

For each momentum mode we can define two new operators,

$$\begin{aligned} \boldsymbol{\alpha^\dagger_k} &\equiv \boldsymbol{\cos\theta_k b^\dagger_{\vec{k}}} + \boldsymbol{\sin\theta_k d_{-\mathrm{k}}}, \\ \boldsymbol{\beta^\dagger_k} &\equiv \boldsymbol{\cos\theta_k d^\dagger_{-\vec{k}}} - \boldsymbol{\sin\theta_k b_{\vec{k}}}. \end{aligned} \tag{13.71}$$

These operators satisfy the commutation relations,

$$\begin{aligned} \boldsymbol{\{\alpha_k, \alpha^\dagger_q\} = \{\beta_k, \beta^\dagger_q\} = \delta_{k,q}}, \\ \boldsymbol{\{\alpha_k, \beta_q\} = \{\beta_k, \alpha^\dagger_q\} = \{\alpha_k, \beta^\dagger_q\} = \{\alpha^\dagger_k, \beta^\dagger_q\} = 0}. \end{aligned} \tag{13.72}$$

If one considers the Hamiltonian,

$$\begin{aligned} \boldsymbol{H} &= \sum_{\boldsymbol{k}} \boldsymbol{E_k} (\boldsymbol{\alpha^\dagger_k \alpha_k} + \boldsymbol{\beta^\dagger_k \beta_k}) \\ &= \sum_{\boldsymbol{k}} \boldsymbol{E_k} \left\{ \boldsymbol{\cos 2\theta_k} \left[\boldsymbol{b^\dagger_{\vec{k}} b_{\vec{k}}} + \boldsymbol{d^\dagger_{-\vec{k}} d_{-\vec{k}}} \right] + \boldsymbol{2\sin^2\theta_k} + \boldsymbol{\sin 2\theta_k} \left[\boldsymbol{b^\dagger_{\vec{k}} d^\dagger_{-\vec{k}}} + \boldsymbol{d_{-\vec{k}} b_{\vec{k}}} \right] \right\}, \end{aligned} \tag{13.73}$$

one may see that the $\boldsymbol{H}$ will equal $\boldsymbol{H_0} + \boldsymbol{V}$ to within a constant if

$$\begin{aligned} \boldsymbol{E_k \cos 2\theta_k} &= \boldsymbol{\hbar c k}, \\ \boldsymbol{E_k \sin 2\theta_k} &= \boldsymbol{mc^2}, \end{aligned} \tag{13.74}$$

or equivalently,

$$\begin{aligned} \boldsymbol{E_k} &= \boldsymbol{\sqrt{\hbar^2 k^2 c^2 + m^2 c^4}}, \\ \boldsymbol{\tan 2\theta_k} &= \boldsymbol{\frac{mc}{\hbar k}}. \end{aligned} \tag{13.75}$$

Thus,

$$\begin{aligned} H &= H_0 + V + 2\sum_k E_k \sin^2\theta_k \qquad (13.76)\\ &= H_0 + V + \sum_k E_k(1-\cos 2\theta_k) \\ &= H_0 + V + \sum_k E_k\left(1 - \sqrt{\frac{1}{1+\tan^2 2\theta_k}}\right) \\ &= H_0 + V + \sum_k \left(\sqrt{(\hbar c k)^2 + m^2c^4} - \hbar c k\right). \end{aligned}$$

The last term is a correction to the vacuum energy which is simply the difference between the vacuum energies with and without the mass. This can be thought of as an correction to the energy of the Dirac sea.

13.12 Exercises

1. To show why derivatives are defined as shown in Eq. (13.13), show that

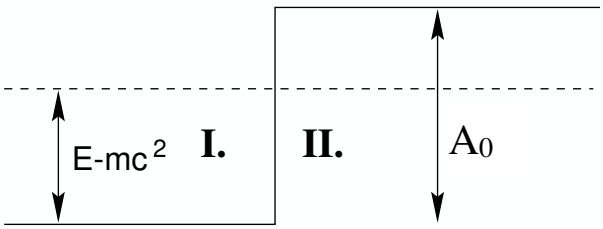
$$\partial_\mu x^2 = 2x_\mu, \text{ and } \partial^\mu x^2 = 2x^\mu,$$

   where $x^2 = x_0^2 - x_1^2 - x_2^2 - x_3^2$.

2. Consider a charged relativistic particle interacting with the electromagnetic field, and described by the Klein-Gordon equation.

$$\left[(i\hbar\partial_t - eA_0)^2 + c^2\hbar^2\partial_x^2 - m^2c^4\right]\psi(x,t) = 0$$

   The electrostatic potential A_0 is illustrated in the diagram below.



   Consider a solution for a particle incident from the left,

$$\begin{aligned} \psi_I(x,t) &= e^{(-iEt+ikx)/\hbar} + Be^{(-iEt-ikx)/\hbar} \\ \psi_{II}(x,t) &= Ce^{(-iEt+ik'x)/\hbar}, \end{aligned}$$

   where $E = \sqrt{m^2c^4 + k^2}$.

   Calculate the charge and current densities (include direction) in regions I and II for each of the following three cases.

   (a) $eA_0 < E - mc^2$.

(b) $E - mc^2 < eA_0 < E + mc^2$.
(c) $eA_0 > E + mc^2$.

3. Consider the same case as above, except with no electrostatic potential. Instead, consider a different mass in region I and region II, with $m_{II} > m_I$. For each of the following two cases, calculate the charge and current densities in regions I and II.

   (a) $E > m_{II}c^2$
   (b) $E < m_{II}c^2$

4. Consider the Dirac representation,

$$\beta = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix} \qquad \vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}$$

   and the chiral representation,

$$\beta = \begin{pmatrix} 0 & -\mathbb{I} \\ -\mathbb{I} & 0 \end{pmatrix} \qquad \vec{\alpha} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & -\vec{\sigma} \end{pmatrix}$$

   The spinors, $u_\uparrow$ and $u_\downarrow$, represent positve-energy eigenvalues of the Dirac equation assuming the momentum is along the z axis.

$$(m\beta + p_z\alpha_z)\, u(p_z) = Eu(p_z)\ ,$$

   The spin labels, $\uparrow$ and $\downarrow$ refer to the positive and negative values of the spin operator,

$$\Sigma_z = \begin{pmatrix} \sigma_z & 0 \\ 0 & \sigma_z \end{pmatrix}$$

   Write the four-component spinors $u_\uparrow$ and $u_\downarrow$ in terms of p, E and m :

   (a) in the Dirac representation.
   (b) in the chiral representation.
   (c) in the limit $p_z \to 0$ for both representations.
   (d) in the limit $p_z \to \infty$ for both representations.

5. Consider a solution to the Dirac equation for massless particles, $u_+(\vec{p})$, where the $+$ denotes the fact that the solution is an eigenstate of the spin operator in the $\hat{p}$ directions,

$$(\vec{\Sigma} \cdot \hat{p}) u_+(\vec{p}, x) = u_+(\vec{p}, x).$$

   Show that the operator β operating on $u_+(p)$ gives a negative energy solution but is still an eigenstate of $\vec{\Sigma} \cdot \hat{p}$ with eigenvalue $+1$.

6. Show that the operator

$$P = \frac{1}{2E_p}(E_p + \vec{\alpha} \cdot \vec{p} + \beta m)$$

(a) is a projector, i.e. $\boldsymbol{P^2 = P}$.

(b) gives a positive-energy solution to the Dirac equation when operating on any state $|\psi\rangle$,

$$(\boldsymbol{E_p - \vec{\alpha}\cdot\vec{p} - \beta m})\boldsymbol{P}|\psi\rangle = \boldsymbol{0}.$$

7. Consider a massless spin half particle of charge $\boldsymbol{e}$ in a magnetic field in the $\hat{z}$ direction described by the vector potential

$$\boldsymbol{\vec{A} = Bx\hat{y}}.$$

The Hamiltonian is then

$$\boldsymbol{H = \alpha_x(-i\hbar\partial_x) + \alpha_y(-i\hbar\partial_y - eBx)}.$$

(a) Show that the Hamiltonian commutes with $\boldsymbol{-i\hbar\partial_y}$ and $\boldsymbol{-i\hbar\partial_z}$.

(b) The wave function can then be written as

$$\boldsymbol{\psi_{k_y,k_z}(x,y,z) = e^{ik_y y + ik_z z}\phi_{k_y,k_z}(x)},$$

After setting $\boldsymbol{k_y = k_z = 0}$, show that the energy can be found by solving the equation

$$\boldsymbol{E^2\phi_\pm(x) = (-\hbar^2\partial_x^2 + e^2B^2x^2 - e\hbar B\Sigma_z)\phi_\pm(x)}.$$

(c) Show that the eigenvalues of the operator $\boldsymbol{H^2}$ are

$$\boldsymbol{E_\pm^2 = (2n + 1 \mp 1)e\hbar B, \;\; n = 0, 1, 2\cdot},$$

where the $\pm$ refers to eigenvalues of Σ_z. You can do this mapping to the harmonic oscillator and then using the solutions to the harmonic oscillator from Chapter 3. Note that when the the eigenvalue of Σ_z is $+1$, there exists a solution with $\boldsymbol{E = 0}$.

8. Using the definition of field operators in Eq. (13.66), show that the Hamiltonian

$$\begin{aligned}\boldsymbol{H} &= \int \boldsymbol{d^3r\; \Psi^\dagger(\vec{r},t)(-i\hbar\vec{\alpha}\cdot\nabla + \beta m)\Psi(\vec{r},t)} \\ &= \sum_{s,\vec{p}} \boldsymbol{E_p(b^\dagger_{s,\vec{p}}b_{s,\vec{p}} + d^\dagger_{s,\vec{p}}d_{s,\vec{p}} - 1)}.\end{aligned} \tag{13.77}$$

I.e. the vacuum energy for each mode is negative.

9. Using the definitions for $\boldsymbol{\alpha_k}$ and $\boldsymbol{\beta_k}$ in Eq. (13.71),

(a) Show that

$$\boldsymbol{b^\dagger_{\vec{k}}b_{\vec{k}} - d^\dagger_{-\vec{k}}d_{-\vec{k}} = \alpha^\dagger\alpha_k - \beta^\dagger\beta_k}.$$

This demonstrates that the eigenstates of the new Hamiltonian are still eigenstates of the charge operator written in the old basis.

(b) Show that the state

$$|\tilde{0}\rangle \equiv \boldsymbol{\cos\theta_k}|0\rangle + \boldsymbol{\sin\theta_k d^\dagger_{-\vec{k}}b^\dagger_{\vec{k}}}|0\rangle$$

is destroyed by both $\boldsymbol{\alpha_k}$ and $\boldsymbol{\beta_k}$, where $|0\rangle$ is the vacuum in the old basis. Effectively, this shows that $|\tilde{0}\rangle$ is the vacuum in the new basis.

14 States Without Conserved Particle Number

Thus far, the initial and final states that we have considered all have a fixed number of particles. That number might have changed between the initial and final state, e.g. an excited state emitting a photon, but the basis being used was purely one where the n particles were well defined, typically by their momentum. However, there are numerous applications where the basis being used is not one of fixed particle number. This is even the case when the particles carry charge. In the next section the Bogoliubov treatment of pairing is presented. In this treatment a variational wave function involving a mixture of states with charge zero (zero electrons) and charge $-2e$ (2 electrons) is proposed. The subsequent section introduces the idea of coherent state for bosons. This basis was applied by Roy Glauber to describe the photon pulse from a laser. It is also serves as a basis for lattice gauge theory. These topics will not be pursued in great detail, but a student should get a decent knowledge of such bases, and perhaps gain some level of comfort with this peculiar choice of states.

14.1 Bogoliubov Operators and Cooper Pairs

The pairing term in BCS theory can be written in terms of creation and destruction operators as

$$V = -\frac{v_0}{\Omega}{\sum_{k,k'}}' b^\dagger_{\vec{k}} d^\dagger_{-\vec{k}} d_{-\vec{k}'} b_{\vec{k}'}. \tag{14.1}$$

Here, we have only considered pairs that sum to total momentum zero. The $b^\dagger$ and $d^\dagger$ operators refer to electron spin-up and spin-down creation operators. We do this rather than fouling the formalism with more subscripts. The interaction represents the scattering of particles of opposite momenta $\vec{k}'$ and $-\vec{k}'$ into states $\vec{k}$ and $-\vec{k}$. The primed sum limits the sum to a region near the Fermi surface.

Now, we consider as a variational wave function a state where each momentum mode is defined as

$$|\Psi\rangle = \prod_k{}' \left(\cos\theta_k + \sin\theta_k b^\dagger_{-\vec{k}} d^\dagger_{\vec{k}}\right)|0\rangle \tag{14.2}$$

When θ_k becomes non-zero this state is no longer an eigenstate of the number operator or electric charge, because the state mixes components with different numbers of electrons. The expectations of the relevant creation/destruction operator combinations are

$$\begin{aligned} \langle\Psi|b^\dagger_{\vec{k}} b_{\vec{k}} + d^\dagger_{-\vec{k}} d_{-\vec{k}}|\Psi\rangle &= 2\sin^2\theta_k, \\ \langle\Psi|b_{\vec{k}} d_{-\vec{k}}|\Psi\rangle &= \frac{1}{2}\sin 2\theta_k. \end{aligned} \tag{14.3}$$

Next, we consider the energy of particles in the primed region,

$$\begin{aligned} \langle\Psi|H_0 + V|\Psi\rangle &= {\sum_k}' \epsilon_k \langle\Psi|b^\dagger_{\vec{k}} b_{\vec{k}} + d^\dagger_{-\vec{k}} d_{-\vec{k}}|\Psi\rangle - \frac{v_0}{\Omega}{\sum_k}' \langle\Psi|b^\dagger_{\vec{k}} d^\dagger_{-\vec{k}}|\Psi\rangle {\sum_{k'}}' \langle\Psi|b_{\vec{k}'} d_{-\vec{k}'}|\Psi\rangle \\ &= {\sum_k}' \left(2\epsilon_k \sin^2\theta_k - \frac{\Delta}{2}\sin 2\theta_k\right), \end{aligned} \tag{14.4}$$

where

$$\Delta \equiv \frac{v_0}{2\Omega}\sum_{k'}{}' \sin 2\theta_{k'} \tag{14.5}$$

The important step here is in writing $\langle\Psi|b^\dagger_{\vec{k}}d^\dagger_{-\vec{k}}d_{-\vec{k}'}b_{\vec{k}'}|\Psi\rangle = \langle\Psi|b^\dagger_{\vec{k}}d^\dagger_{-\vec{k}}|\Psi\rangle\langle\Psi|b_{\vec{k}'}d_{-\vec{k}'}|\Psi\rangle$. This follow from the ansatz for the product form of the state, Eq. (14.2), and from ignoring terms where $\vec{k} = \vec{k}'$. These terms are ignored because they are an infinitesimal fraction of the terms in the continuum limit where many states are included in the sum.

Now, one must choose the angles θ_k so that the energy is minimized compared to the extra particles coming from a reservoir with chemical potential μ. (Usually μ is the Fermi energy).

$$\frac{\partial}{\partial\theta_k}\sum_k{}'\left(2(\epsilon_k-\mu)\sin^2\theta_k - \frac{v_0}{4\Omega}\sin 2\theta_k\sum_{k'}{}'\sin 2\theta_{\vec{k}'}\right) = 0. \tag{14.6}$$

This leads to the solution

$$\tan 2\theta_k = \frac{\Delta}{\epsilon_k-\mu}. \tag{14.7}$$

After finding the angles θ_k, one must find the new Δ using Eq. (14.5), then iterate until a consistent solution is found.

One can find the probability that a state $\vec{k}$ is occupied by a particle,

$$\langle b^\dagger_{\vec{k}}b_{\vec{k}}\rangle = \sin^2\theta_{\vec{k}} = \frac{1}{2} - \frac{1}{2}\frac{(\epsilon_k-\mu)}{\sqrt{(\epsilon_k-\mu)^2+\Delta^2}}.$$

The occupation is 1/2 when $\epsilon_k = \mu$, is zero for $\epsilon_k >> \mu$ and is unity for $\epsilon_k << \mu$. Pairing smooths the step function one expects for non-interacting particles with a scale Δ.

We also note the difference of this solution with the relativistic Dirac problem above. In that case the ground state was a mixture of the old vacuum and a particle/anti-particle pair, i.e. $|0\rangle$ and $b^\dagger_k d^\dagger_{-k}|0\rangle$. However, in this approach the solution is a mixture of zero-electron and two-electron states, thus not even electric charge is fixed. The analogy between superconductivity and particle/nuclear physics is reflected in the jargon, as the generation of masses due to spontaneous symmetry breaking is referred to as the creation of a gap.

Example 14.1: Show that same-spin particles don't pair
If we had considered particles of the same spin,

$$|\Psi\rangle = \prod_k{}'\left(\cos\theta_k + \sin\theta_k b^\dagger_{\rm k}b^\dagger_{-\vec{k}}\right)|0_k\rangle,$$

the resulting expectation for the pairing would be zero. This follows because

$$(b_{\vec{k}}b_{-\vec{k}} + b_{-\vec{k}}b_{\vec{k}})\left(\cos\theta_k + \sin\theta_k b^\dagger_{\vec{k}'}b^\dagger_{-\vec{k}'}\right)|0\rangle = 0.$$

The cancelation results from the fact that $\vec{k}$ could equal either $\vec{k}'$ or $-\vec{k}'$. There is only one possible connection for the case where the product of $b_{\vec{k}}b_{-\rm k}$ is replaced by $b_{\vec{k}}d_{-\rm k}$. One way

of stating this result is that an s-wave coupling must combine with an anti-symmetric spin combination because the overall pair wave function must be anti-symmetric. If the original potential had odd-parity, e.g. it was proportional to k, then one could form a p wave like state with symmetric spin. The wave function might have a form

$$|\Psi\rangle = \prod_k{}' \left(\cos\theta_k + f(\vec{k})\sin\theta_k b^\dagger_{\vec{k}} b^\dagger_{-\vec{k}}\right)|0_k\rangle,$$

where f is an odd-parity function of $\vec{k}$.

14.2 Coherent States

Our last example of a state which is not an eigenstate of particle number is referred to as a "coherent" state.

$$\begin{aligned} |\eta\rangle &= e^{-\eta^*\eta/2} e^{\eta a^\dagger}|0\rangle \\ &= e^{-\eta^*\eta/2}\sum_n \frac{(\eta a^\dagger)^n}{n!}|0\rangle, \end{aligned} \tag{14.8}$$

where η is a complex number. One can check to see that this state is properly normalized by calculating the overlap,

$$\begin{aligned} \langle 0|e^{\eta^* a} e^{\eta a^\dagger}|0\rangle &= \langle 0|\sum_m \frac{(\eta^* a)^m}{m!}\sum_n \frac{(\eta a^\dagger)^n}{n!}|0\rangle \\ &= \sum_n \langle 0|\frac{(\eta^*\eta)^n}{(n!)^2} a^n (a^\dagger)^n|0\rangle \\ &= \sum_n \frac{(\eta^*\eta)^n}{n!} = e^{\eta^*\eta}. \end{aligned} \tag{14.9}$$

A coherent state is an eigenstate of the destruction operator.

$$\begin{aligned} a|\eta\rangle &= e^{-\eta^*\eta/2} a\sum_n \frac{(\eta a^\dagger)^n}{n!}|0\rangle \\ &= \sum_n \frac{n\eta^n (a^\dagger)^{n-1}}{n!}|0\rangle \\ &= \eta\sum_n \frac{\eta^{n-1}(a^\dagger)^{n-1}}{(n-1)!}|0\rangle \\ &= \eta|\eta\rangle. \end{aligned} \tag{14.10}$$

Thus, it is simple to calculate matrix elements of coherent states. For example,

$$\langle\gamma|(a^\dagger)^3 a^2|\eta\rangle = (\gamma^*)^3\eta^2\langle\gamma|\eta\rangle. \tag{14.11}$$

Finally, the same coherent state in Eq. (14.8) can be expressed as

$$|\eta\rangle = e^{-\eta^*\eta/2} e^{\eta a^\dagger}|0\rangle = e^{\eta a^\dagger - \eta^* a}|0\rangle. \tag{14.12}$$

Because the argument of the exponential in the latter form is manifestly anti-Hermitian, it is easy to see that $\langle\eta|\eta\rangle = 1$. It is left as part of a homework problem to prove the equivalence of the two forms (use the Baker-Campbell-Hausdorff theorem).

14.3 Emission Via Coupling to a Classical Current

The physical importance of coherent states comes from the fact that they are the solution to the following Hamiltonian.

$$\begin{aligned} H(t) &= H_0 + V(t). \\ H_0 &= \epsilon a^\dagger a, \\ V(t) &= j(t)[a^\dagger + a]. \end{aligned} \tag{14.13}$$

Here $j(t)$ plays the role of an external classical current that couples to a quantum field, e.g. $j \cdot A$.

As a solution of the Hamiltonian, one can propose a state $|\eta(t)\rangle_I$ in the interaction representation,

$$\begin{aligned} |\eta(t)\rangle_I = e^{-\eta^*(t)\eta(t)/2} e^{\eta(t)a^\dagger}|0\rangle = e^{\eta(t)a^\dagger - \eta^*(t)a}|0\rangle, \\ \text{where } \eta(t) \equiv \frac{-i}{\hbar}\int_{-\infty}^{t} dt' e^{i\epsilon t'/\hbar} j(t'). \end{aligned} \tag{14.14}$$

This state is manifestly normalized because the argument of the exponential is anti-Hermitian. Viewing the evolution of $|\eta\rangle_I$ with time,

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}|\eta(t)\rangle_I = \left(j(t)e^{i\epsilon t/\hbar}a^\dagger + j(t)e^{-i\epsilon t/\hbar}a\right)|\eta(t)\rangle_I \\ j(t)\left(a^\dagger(t) + a(t)\right))|\eta(t)\rangle_I. \end{aligned} \tag{14.15}$$

Here, $a(t)$ is the destruction operator in the interaction representation,

$$\begin{aligned} a(t) &= e^{-iH_0 t/\hbar} a e^{iH_0 t/\hbar} = e^{-i\epsilon t/\hbar} a, \\ a^\dagger(t) &= e^{-iH_0 t/\hbar} a^\dagger e^{iH_0 t/\hbar} = e^{i\epsilon t/\hbar} a^\dagger. \end{aligned} \tag{14.16}$$

This then gives

$$i\hbar\frac{\partial}{\partial t}|\eta(t)\rangle_I = V_I(t)|\eta(t)\rangle_I, \tag{14.17}$$

which demonstrates that our guess at a form for $|\eta(t)\rangle_I$ was successful.

Coupling to an external current is a common consideration in a wide variety of physical problems. Even when spatial degrees of freedom are considered, e.g. $j \cdot A$, one finds simple analytic solutions. They then often have the form

$$|\eta\rangle \sim \exp\left\{-i\int d^3p\, j(\vec{p}) a^\dagger_{\vec{p}}\right\}|0\rangle, \quad j(\vec{p}) \sim \int d^4x\, e^{ip\cdot x} j(x). \tag{14.18}$$

Any classical current that couples linearly to the field will radiate particles in a coherent state. Some lasers are described with coherent states, when one can justify that the current behaves classically. A classical current is one that is unaffected by the emission. When an atom emits a photon the atom then changes to a new state. For a classical current, no such change occurs. To justify treating a current as classical, one typically considers macroscopic emitters. For a laser, one can think of a crystal which is coupled to the field as providing the classical current.

14.4 Completeness Relations for Coherent States

A useful property of any basis is completeness,

$$\sum_{\alpha} \langle m|\alpha\rangle\langle\alpha|n\rangle = \delta_{mn}. \tag{14.19}$$

Coherent states provide a complete basis when all values of η in the complex plane are considered.

$$\frac{1}{\pi}\int d\eta_R d\eta_I \langle m|\eta\rangle\langle\eta|n\rangle = \delta_{mn}, \tag{14.20}$$

where η_R and η_I are the real and imaginary parts of η. To prove this, we expand the matrix elements.

$$\begin{aligned}\langle m|\eta\rangle &= e^{-|\eta|^2/2}\langle m|\frac{(\eta a^\dagger)^m}{m!}|0\rangle, \\ &= e^{-|\eta|^2/2}\frac{\eta^m}{\sqrt{m!}}, \\ \langle m|\eta\rangle\langle\eta|n\rangle &= e^{-|\eta|^2}\frac{\eta^m(\eta^*)^n}{\sqrt{n!m!}}.\end{aligned} \tag{14.21}$$

Writing the integral over the real and imaginary parts of η as

$$\int d\eta_R d\eta_I \to \int d\phi|\eta|d|\eta|, \tag{14.22}$$

where ϕ is the complex phase of η, allows one to see that the integral over ϕ will eliminate all terms with $m \neq n$ because

$$\begin{aligned}\int d\phi e^{i(m-n)\phi} &= 2\pi\delta_{mn}, \\ \langle m|\eta\rangle\langle\eta|n\rangle &= e^{-|\eta|^2}e^{i(m-n)\phi}\frac{|\eta|^{m+n}}{\sqrt{n!m!}}.\end{aligned} \tag{14.23}$$

One may now rewrite the expression using the fact that integrating over ϕ constrains the result to be proportional to δ_{mn}.

$$\frac{1}{\pi}\int d\eta_R d\eta_I \langle m|\eta\rangle\langle\eta|n\rangle = 2\delta_{mn}\int |\eta|d|\eta|\; e^{-|\eta|^2}\frac{|\eta|^{2n}}{n!}. \tag{14.24}$$

By making the substitution $u \equiv |\eta|^2$, one can perform the integral and see that

$$\frac{1}{\pi}\int d\eta_R d\eta_I \langle m|\eta\rangle\langle\eta|n\rangle = \delta_{mn}. \tag{14.25}$$

14.5 Path Integrals

Path integrals provide an alternative means to express quantum mechanics. Rather than discrete states, the basis for each mode becomes coherent states described by some complex number η. Path integrals form the basis for some non-perturbative methods, principally lattice gauge theory.

Completeness as displayed in the previous section, combined with the fact that the coherent states are eigenstates of the destruction operator, allows one to calculate all matrix elements

$$\begin{aligned}\langle \alpha_f | e^{-iHt/\hbar} | \alpha_i \rangle = \sum_{\alpha_1, \alpha_2 \cdots} & \langle \alpha_f | e^{-iH(t-t_1)/\hbar} | \alpha_1 \rangle \langle \alpha_1 | e^{-iH(t_1-t_2)/\hbar} | \alpha_2 \rangle \\ & \langle \alpha_2 | e^{-iH(t_2-t_3)/\hbar} | \alpha_3 \rangle \cdots \langle \alpha_n | e^{-iH(t_n-t_i)/\hbar} | \alpha_i \rangle\end{aligned} \tag{14.26}$$

as an integral over complex fields η_n rather than a sum over discrete states α_n. Furthermore, if the Hamiltonian is normal ordered, that is each term has all creation operators pushed to the left,

$$\langle \eta_n | e^{-iH(a^\dagger, a)\delta t/\hbar} | \eta_{n+1} \rangle = \left\{ 1 - iH(\eta_n^*, \eta_{n+1})\delta t/\hbar \right\} \langle \eta_n | \eta_{n+1} \rangle . \tag{14.27}$$

Going further,

$$\begin{aligned}\langle \eta_n | \eta_{n+1} \rangle &= e^{-\eta^* \eta/2} \langle 0 | e^{\eta a} | \eta + \delta\eta \rangle \\ &= e^{-\eta^* \eta/2} e^{\eta^* (\eta + \delta\eta)} \langle 0 | \eta + \delta\eta \rangle \\ &= e^{-\eta^* \eta/2} e^{\eta^* (\eta + \delta\eta)} e^{(\eta + \delta\eta)^* (\eta + \delta\eta)/2} \\ &= e^{(\eta^* \delta\eta - \delta\eta^* \eta)/2} \\ &= e^{(\eta^* \dot{\eta} - \dot{\eta}^* \eta)\delta t/2} .\end{aligned} \tag{14.28}$$

This gives

$$\langle \eta_n | e^{-iH(a^\dagger, a)\delta t/\hbar} | \eta_{n+1} \rangle = 1 + \left[-iH(\eta_n^*, \eta_{n+1})/\hbar + (\eta^* \dot{\eta} - \dot{\eta}^* \eta)/2 \right] \delta t. \tag{14.29}$$

Thus, the matrix element is transformed into a purely numerical function of the complex number η. The process of breaking up the evolution operator into the product of many individual pieces $e^{-iH\delta t/\hbar}$, inserting a complete set of states between each piece, expressing them in terms of some variables at each point (η) and integrating over all η is known as a path integral.

Finally, we point out that rather than thinking of the variable η, one can also write the integral in terms of the real variables p and q, where $\eta = (p + iq)/\sqrt{2\hbar}$. In that case

$$\frac{1}{\pi} d\eta_R d\eta_I \rightarrow \frac{1}{2\pi\hbar} dp dq \, , \tag{14.30}$$

and

$$\langle \eta_n | e^{-iH(a^\dagger, a)\delta t/\hbar} | \eta_{n+1} \rangle = \{ 1 + i[p\dot{q}/2 - q\dot{p}/2 - H(p, q)]\delta t/\hbar \} \, . \tag{14.31}$$

The final matrix element is thus

$$\langle \eta(t_f) | \eta(t_0) \rangle = \Pi_i \frac{dp(t_i) dq(t_i)}{2\pi\hbar} \exp \left\{ i \int dt [(p\dot{q} - q\dot{p})/2 - H(p, q)]/\hbar \right\} . \tag{14.32}$$

Because one integrates over all t, the contribution from a term $e^{i\int dt(p\dot{q}+q\dot{p})/\hbar}$ is independent of the intermediate values of p and q, as long as one is looking at specific initial and final states. I.e,

$$\begin{aligned}\int_{t_i}^{t_f} dt[\dot{p}q + p\dot{q}] &= \int_{t_i}^{t_f} dt(d/dt)(pq) \\ &= p(t_f)q(t_f) - p(t_i)q(t_i).\end{aligned} \tag{14.33}$$

Thus, it leads to a path-independent phase and can be neglected. This allows one to make the replacement in Eq. (14.31),

$$i(p\dot{q} - q\dot{p})/2\hbar \to ip\dot{q}/\hbar. \tag{14.34}$$

The final matrix element becomes

$$\langle \eta(t_f)|\eta(t_0)\rangle = \Pi_i \frac{dp(t_i)dq(t_i)}{2\pi\hbar} \exp\left\{ i\int dt[p\dot{q} - H(p,q)]/\hbar \right\}. \tag{14.35}$$

Lo and behold, this phase looks like the Lagrangian. Thus, quantum physics can be re-expressed as a path integral over coordinates p and q. In quantum field theory, every point in space is reduced to a state, using such coordinates. This is how Lagrangian field theory, which is the theoretical backbone of the standard model, is described in terms of fields at each point in space encapsulated by numbers, can be translated into quantum mechanics.

In the limit that $\hbar \to 0$, the only path that should contribute is one where the phase,

$$S \equiv \int dt\ [p\dot{q} - H(p,q)] = \int dt\ L(p,q), \tag{14.36}$$

is stationary with respect to small variations of the path. The quantity S is known as the action, and using the calculus of variations leads to Lagrange's classical equations of motion.

In addition to being occasionally useful for calculating evolution operators, path integrals can be applied to thermodynamic partition functions which require calculating the trace of $e^{-\beta H}$. Real-time path integrals tend to be problematic because of the complex phase factors. Because the integrals are typically performed by Monte Carlo procedures there is a huge amount of cancellation from different paths, which makes the calculations extremely noisy. Because the thermodynamic trace considers a path in imaginary time, the calculations are far less noisy. Lattice gauge theory is the numeric enterprise of calculating such path integrals, and is most often employed for QCD.

14.6 Exercises

1. Consider the coherent state $|\eta\rangle$ defined by,

$$|\eta\rangle = e^{-\eta^*\eta/2} \exp\left(\eta a^\dagger\right)|0\rangle.$$

   (a) Show that $|\eta\rangle$ can also be written as

$$|\eta\rangle = e^{-\eta^* a + \eta a^\dagger}|0\rangle.$$

   Hint: You may wish to use the Baker-Campbell-Hausdorff lemma.

(b) Show that the overlap of two states is given by,

$$\langle \eta' | \eta \rangle = e^{-|\eta'|^2/2 - |\eta|^2/2 + \eta'^* \eta}.$$

2. Consider bosonic creation and destruction operators, $a^\dagger$ and a. Consider a linear combination,

$$b = \alpha a + \beta a^\dagger$$

What is the constraint on the complex numbers α and β if one is to demand that $[b, b^\dagger] = 1$?

3. Consider a coherent state

$$|\eta\rangle = e^{-\eta^* \eta/2} e^{\eta a^\dagger} |0\rangle.$$

(a) Show that $\bar{N} = \langle \eta | N_{\rm op} | \eta \rangle = \eta^* \eta$, where $N_{\rm op} = a^\dagger a$ is the number operator.

(b) Show that the variance equals the mean, i.e.,

$$\langle \eta | (N_{\rm op} - \bar{N})^2 | \eta \rangle = \bar{N}.$$

This is characteristic of a Poissonian distribution.