Contents

1 Formal Quantum Mechanics

1.1 Postulates of Quantum Mechanics

We will start with a list of postulates of quantum mechanics. This list is intentionally not numbered, as the postulates are not axioms, and should not be taken as such. Perhaps some are derivable using others, but the following list should be somewhat self-consistent.

Firtly, all mechanical properties of a mechanical system are obtainable from its state function. That is any system, single or multiple particle, with forces or without them, any system is from our perspective reducable to a single, complex valued scalar state function.

The state function is often expressed as wavefunction, which obeys Born rule, i.e.

$$\int_V |\Psi(\vec{r}, t)| dV = 1 \tag{1}$$

where V is the configuration space, $(\vec{r}, t)$ is a vector in the configuration space and dV is the element of the configuration space. Often, we also say that such a function is normalised.

The mechanical properties obtainable from the state function that are experimentally accesable are called observables. They are obtained via their associated operators, which usually take some differential operator form. These operators transform the function into another function, which can be then further used to find the value of the observable. The observable values are always real. Operators are usually denoted as $\hat{Q}$, where Q is the classical symbol for the observable.

When a measurement is made, the resulting value of the observable is an eigenvalue of the associated operator acting on a function. If the state function is not the eigenfunction of the operator, linear superposition of eigenfunctions of the operator is used to express the state function. After the measurement, the state function becomes the eigenfunction of the operator of the observable measured. If there are multiple eigenvalues possible (original function was not an eigenfunction), the eigenvalue and new eigenfunction are randomly with probability $|c_n|^2$, where c_n is the complex coefficient in the superposition series for the old function.

The expectation value of an observable $< Q >$ (average value after many measurements under exactly the same conditions) is given by

$$< Q > = \int_V \Psi^* \hat{Q} \Psi dV \tag{2}$$

States that are not being measured evolve according to time dependent Schrödinger equation

$$\hat{H}\Psi = i\hbar \frac{\partial}{\partial t}\Psi \tag{3}$$

where $\hat{H}$ is the Hamiltonian total energy operator.

Lastly, forms of operators are either derived from their inner relationships or using the correspondence principle - operators are build from a set of two basic operators - position and momentum, or position and energy - by writing the classical expression in terms of position and momentum etc. and then replacing the variables by operators.

The postulates on their own are rather abstract, they are best only referenced in time when needed, otherwise ideally remembered as a whole.

1.1.1 Principle of Superposition

Say we have a general wavefunction Ψ and want to operate on it with an operator $\hat{Q}$, which has eigenfunctions Ψ_n and their corresponding eigenvalues q_n. Then we represent the function Ψ as

$$\Psi = \sum_n c_n \Psi_n \tag{4}$$

Therefore, the expectation value of the variable Q is (m runs over the range of n)

$$< Q > = \int_V \Psi^* \hat{Q} \Psi dV = \int_V \left(\sum_m c_m \Psi_m\right)^* \hat{Q} \left(\sum_n c_n \Psi_n\right) dV = \int_V \left(\sum_m c_m^* \Psi_m^*\right)\left(\sum_n c_n \hat{Q} \Psi_n\right) dV =$$

$$= \int_V \left(\sum_m c_m^* \Psi_m^*\right)(c_n q_n \Psi_n) dV = \int_V \left(\sum_n q_n \sum_m c_n c_m^* \Psi_m^* \Psi_n\right) dV = \sum_n q_n \sum_m c_n c_m^* \int_V \Psi_m^* \Psi_n dV$$

The expectation value is then some form of weighted sum of eigenvalues of all eigenfunctions of the operator. Further advance can be made if the eigenfunctions are orthonormal.

1.1.2   Orthonormal Eigenfunctions

We say that the eigenfunctions are orthonormal iff

$$\int_V \Psi_m^* \Psi_n = \delta_{mn} \tag{5}$$

where δ_{mn} is the Kronecker delta.
Then, the expectation value simplifies to

$$< Q > = \sum_n q_n \sum_m c_n c_m^* \int_V \Psi_m^* \Psi_n dV = \sum_n q_n \sum_m c_n c_m^* \delta_{nm} = \sum_n q_n |c_n|^2 \tag{6}$$

This corresponds to our macroscopical perspective - each measurement measures q_n with probability $|c_n|^2$ - the expectation value is then naturally the weighted sum of these values by their probabilities.

$$< Q > = \sum_n q_n P_n$$

which we derived from the orthonormality condition and our postulates.
One more observation can be made - we required that Ψ is normalized. This means that

$$\int_V \Psi^* \Psi dV = 1$$

$$1 = \int_V \left(\sum_m c_m^* \Psi_m^* \right) \left(\sum_n c_n \Psi_n \right) dV = \sum_n \sum_m c_n c_m^* \int_V \Psi_m^* \Psi_n dV = \sum_n \sum_m c_n c_m^* \delta_{nm} = \sum_n |c_n|^2$$

Hence

$$\sum_n |c_n|^2 = 1 \tag{7}$$

1.1.3   Hilbert space

We can notice that the eigenfunctions of an operator form a Hilbert space of functions, with inner product defined as

$$< \Psi_m, \Psi_n > = \int_V \Psi_m^* \Psi_n dV$$

This operation is positive definite for any eigenfunction with itself.
All requirements can be checked, but we will not discuss them here. However, it should be noted that the consequences of the functions being part of a Hilbert space with given inner product follow, so for example Cauchy-Schwarz inequality holds

$$< f, f >< g, g > \geq | < f, g > |^2 = | < g, f > |^2$$

In our language, this means that

$$\left(\int_V f^* f dV \right) \left(\int_V g^* g dV \right) \geq \left| \int_V f^* g dV \right|^2 = \left| \int_V g^* f dV \right|^2 \tag{8}$$

1.2   Dirac Notation

Dirac notation is a notation for the inner products in Hilbert space modified for Quantum mechanics. The inner product is denoted as

$$< f, g > = < f | g >$$

and if we search for an expectation value, it is denoted as

$$< Q > = \int_V f^* \hat{Q} g dV = < f | \hat{Q} | g >$$

Furthemore, the second vector/function $|g>$ is used to represent the function g itself. We often do not write the function g inside these so called "kets" explicitly, rather we write some label, such as $|up> = g$ or $|1> = g$.

The other part of the expression is called the "bra" part (together form a "bracket"). This expression is essentially a integral operator expression

$$< f| = \int_V dV f^*$$

Many expressions and relations we will derive have more managable form in the Dirac notation, and it is especially useful to denote states for which we do not now the wavefunction explicitly.
Some of the expressions we already derived are for example Cauchy-Schwarz inequality

$$< f|f >< g|g >\geq | < f|g > |^2$$

Normalization condition

$$< \Psi|\Psi >= 1$$

Orthonormality condition

$$< \Psi_m|\Psi_n >= \delta_{nm}$$

and we could go on and on.

1.3   Correspondence Principle

In order to be able to derive explicit forms of operators, we need to choose at least two specific operators. First we choose is the position operator, $\hat{r}$. We choose the wavefunction to exist in real space, and then require the position operator to be simply multiplicative

$$\hat{r}\Psi = \vec{r}\Psi$$

This means that the wavefunction will be required to somehow represent the position of the particle, which is intuitively understandable. This also means that when measuring the position, the posibility of observing a particle at certain position is proportional only to the modulus squared of Ψ. We then say that $|\Psi|^2$ is the probability density of observing particle at certain position.
We choose the momentum operator with analogy to plane waves. Lets imagine that Ψ is a right travelling wave. Then

$$\Psi = \Psi_0 e^{i(\vec{k}\cdot\vec{r}-\omega t)}$$

$$\nabla\Psi = \Psi_0 e^{i(\vec{k}\cdot\vec{r}-\omega t)} \left(ik_x\hat{i} + ik_y\hat{j} + ik_z\hat{k}\right) = i\vec{k}\Psi$$

Hence

$$-i\hbar\nabla\Psi = \hbar\vec{k}\Psi$$

By de Broglie relation

$$-i\hbar\nabla\Psi = \vec{p}\Psi$$

Hence we have our momentum operator

$$\hat{p} = -i\hbar\nabla$$

But, if we choose $\Psi = \Psi_0 e^{-i(\vec{k}\cdot\vec{r}-\omega t)}$, we would end up with $\hat{p} = i\hbar\nabla$. Therefore, the sign before the i is a little bit ambivalent, and depends on preffered notation for plane waves.
From these two operators, we can build other operators. First one that comes to mind is the kinetic energy operator. By correspondence, kinetic energy T is

$$T = \frac{\vec{p}\cdot\vec{p}}{2m} = \frac{|\vec{p}|^2}{2m}$$

Hence in operator language

$$\hat{T} = \frac{\hat{p}\cdot\hat{p}}{2m} = -\hbar^2\frac{\nabla\cdot\nabla}{2m} = -\frac{\hbar^2}{2m}\nabla^2$$

Where ∇^2 is the Laplacian operator.
The potential energy operator can be set to be also only a multiplicative operator, i. e.

$$\hat{V} = V$$

Therefore, we have our total energy hamiltonian operator

$$\hat{H} = \hat{T} + \hat{V} = -\frac{\hbar^2}{2m}\nabla^2 + V$$

5

And TDSE is

$$-\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi = i\hbar\frac{\partial}{\partial t}\Psi$$

Similarly, we can derive forms of operators for angular momentum. In classical physics

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

for point particle. This is a set of three equations

$$L_x = yp_z - zp_y$$

$$L_y = zp_x - xp_z$$

$$L_z = xp_y - yp_x$$

This can be directly translated to operator language

$$\hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y$$

$$\hat{L}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z$$

$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x$$

Here, the $\hat{p}_x$ operator is operator that extracts the x component of the momentum, which is simply the x component of the operator $\hat{p}$, so

$$\hat{p}_x = -i\hbar\frac{\partial}{\partial x}$$

Similarly for other components, and of course $\hat{x} = x$, as for other components as well. The reason these are not written as a vector will become apparent later, but it is essentially because separate components of the angular momentum are not compatible - they cannot be all known at the same time.

1.4   Schrödinger's equations

The states that are not being observed evolve according to time dependent Schrödinger equation. This equation involves a time derivative of the wavefunction, and therefore enables us to describe the time dependence of the wavefunction. Without derivation, the equation is

$$\hat{H}\Psi = i\hbar\frac{\partial}{\partial t}\Psi \tag{9}$$

We know the form of Hamiltonian operator, so the time dependent Schrödinger equation (TDSE) becomes

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\Psi = i\hbar\frac{\partial}{\partial t}\Psi \tag{10}$$

Now, if the potential energy does not change with time, there exists a special stationary class of solutions, for which we can write $\Psi = \psi(\vec{r})\phi(t)$, where $\vec{r}$ is the position vector. Then, the equation becomes

$$-\frac{\hbar^2}{2m}\phi\nabla^2\psi + V\psi\phi = i\hbar\psi\frac{\partial\phi}{\partial t}$$

Dividing by $\psi\phi$

$$-\frac{\hbar^2}{2m}\frac{1}{\psi}\nabla^2\psi + V = i\hbar\frac{1}{\phi}\frac{d\phi}{dt}$$

The left side of the equation is now only the function of position, while the right side is only function of time. Therefore, both sides together must be equal to a constant. Also, as ϕ is only function of time, we can change the partial derivative to total derivative. Then, the equation for time part becomes (naming separation constant E)

$$E = i\hbar\frac{1}{\phi}\frac{d\phi}{dt}$$

$$\frac{d\phi}{dt} = -\frac{iE}{\hbar}\phi$$

This is a simple differential equation with solution

$$\phi = A e^{-i\frac{E}{\hbar}t} \tag{11}$$

But, this means that the total wavefunction only rotates in phase with time, but otherwise does not change, because

$$\Psi = A\psi e^{-i\frac{E}{\hbar}t}$$

Also, the probability density of finding the particle becomes

$$|\Psi|^2 = AA^*|\psi|^2$$

which is completly stationary. Therefore, the only function we need to derive in order to describe stationary state solutions of Schrödinger equation is the ψ function. Using the separation

$$-\frac{\hbar^2}{2m}\frac{1}{\psi}\nabla^2\psi + V = E$$

And therefore

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi = E\psi \tag{12}$$

this is called the time independent Schrödinger equation (TISE), and we can now interpret the constant of separation E as the total energy of the system. The solution to this equation depends on exact form of potential V, and we will discover specific solutions later on. This equation is generally easier to solve than TDSE.

1.5   Models and Matrices

In many cases, we know either from experiment or from other physic's requirements that certain system has only limited amount of states that it can be in. However, these states might not be the eigenstates of the Hamiltonian operator. This means that if Ψ_i is the state we know exists for the system, the operation

$$\hat{H}\Psi_i = \left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\Psi_i \neq E_i(t)\Psi_i = i\hbar\frac{\partial}{\partial t}\Psi_i$$

where the energy of the state E_i can generally depend on time. This means that such states do not obey TDSE. However, we can try to find the energy eigenstates as linear combinations of the known states Ψ_i. Let Ψ' be some eigenstate of hamiltonian operator (for which TDSE applies), and let

$$\Psi' = \sum_i c_i(t)\Psi_i$$

where we allow for the coefficients to depend on time, but not on space (the superposition would otherwise not really make sense).

Lets now, for the sake of brevity, change to Dirac notation. Therefore, the equation above becomes

$$|\Psi'> = \sum_i c_i(t)|\Psi_i>$$

The TDSE is (in Dirac notation)

$$\hat{H}|\Psi'> = i\hbar\frac{\partial}{\partial t}|\Psi'> = E'(t)|\Psi'>$$

where E' is the total energy of state $|\Psi'>$ since this state is an eigenstate. Then, doing integration by $<\Psi_i|$, which is only across the space, we have

$$<\Psi_i|\hat{H}|\Psi'> = E'(t)<\Psi_i|\Psi'>$$

since E' is a constant in space. Writing $|\Psi'>$ in terms of different $|\Psi_i>$ and changing the index in these to j, the equation becomes

$$<\Psi_i|\hat{H}\left(\sum_j c_j|\Psi_j>\right) = E'<\Psi_i|\left(\sum_j c_j\Psi_j\right)$$

Since c_j is simply a constant in space

$$\sum_j c_j < \Psi_i|\hat{H}|\Psi_j >= E' \sum_j c_j < \Psi_i|\Psi_j >$$

Now, if we assume that the states Ψ_i are orthonormal, we have

$$\sum_j < \Psi_i|\hat{H}|\Psi_j > c_j = E' c_i \tag{13}$$

we could write this equation for any i. Therefore, we would have a set of equations, one equation for each state $|\Psi_i >$. The equations would always have $E' c_i$ on the right-hand side and a linear combination over all c_i on the left hand side (the sum over j runs over all i - only a dummy variable).
This relationship can be expressed in form of a matrix. To make this more obvious, we can set

$$H_{ij} =< \Psi_i|\hat{H}|\Psi_j >$$

Then the equation is

$$\sum_j H_{ij} c_j = E' c_i$$

Hence

$$\mathbf{H}\vec{c} = E'\vec{c} = i\hbar \frac{\partial}{\partial t}\vec{c} \tag{14}$$

where $\vec{c}$ is the column vector of coefficients c_i and $\mathbf{H}$ is the matrix $H_{ij} =< \Psi_i|\hat{H}|\Psi_j >$. This means that the vector of coefficients $\vec{c}$ is an eigenvector of matrix $\mathbf{H}$.
We then see why we called state $|\Psi' >$ an eigenstate - all we need for its creation is vector $\vec{c}$, which is an eigenvector of $\mathbf{H}$.
This formalims has great advantage, because $\mathbf{H}$ is all we need to find the eigenvectors, and it does not depend on specific eigenstate we choose - we can always write down each element $H_{ij} =< \Psi_i|\hat{H}|\Psi_j >$ without knowledge of any of eigenstates Ψ'.
These eigenstates rotate each with period $\frac{h}{E}$, which can be than used to determine how the observable states Ψ_i change - it is simply a matter of change of base and expression of Ψ_i using only different vectors $\vec{c}$ for different eigenstates.
This formalism is later on applied to some systems, although generally, it is a bit problematic, as it requires that the matrix is finite. We can get around this if we assume that most of the infinite matrix elements are almost zero, as will be done in almost free electron model later on. Now, however, we will continue with more rigorous approaches, which do not require finite number of states.

1.6   Reality and Hermiticity

In order for the observable to be real, its expectation value needs to be real (all measurements yield real numbers). Let $\hat{Q}$ be an operator of a real observable, and Ψ_n be its eigenfunction. General state of the system Ψ can be described as linear combination of eigenfunctions of $\hat{Q}$ as $\Psi = \sum_n c_n \Psi_n$. The expectation value is

$$< Q >= \int_V \Psi^* \hat{Q} \Psi dV =< \Psi|\hat{Q}|\Psi >$$

Substituting in for Ψ in terms of the eigenfunctions of $\hat{Q}$

$$< Q >= \int_V \left(\sum_m c_m^* \Psi_m^*\right) \hat{Q} \left(\sum_n c_n \Psi_n\right) dV = \sum_m \sum_n c_m^* c_n \int_V \Psi_m^* \hat{Q} \Psi_n dV$$

Or, if we simply rename the indices

$$< Q >= \sum_p \sum_q c_p^* c_q \int_V \Psi_p^* \hat{Q} \Psi_q dV$$

Taking the complex conjugate of the integral (assuming real space element dV) for real $< Q >$

$$< Q >^* =< Q >= \sum_p \sum_q c_p c_q^* \int_V \Psi_p (\hat{Q}\Psi_q)^* dV$$

Now, we define complex conjugate of an operator as

$$\hat{Q}^*\Psi^* = (\hat{Q}\Psi)^* \tag{15}$$

for any Ψ. This is further discussed in following subsection.

Using the complex conjugate of an operator and equating with original expression for $< Q >$

$$< Q >= \sum_p \sum_q c_p c_q^* \int_V \Psi_p \hat{Q}^* \Psi_q^* dV = \sum_m \sum_n c_m^* c_n \int_V \Psi_m^* \hat{Q} \Psi_n dV$$

This can be made equal only if terms with equal constants in front of integrals are equal. This means that for some m and n on the right hand side, we choose a term with $p = n$ and $q = m$, which has to be equal. This means

$$c_n c_m^* \int_V \Psi_n \hat{Q}^* \Psi_m^* dV = c_m^* c_n \int_V \Psi_m^* \hat{Q} \Psi_n dV$$

$$\int_V \hat{Q}^* \Psi_m^* \Psi_n dV = \int_V \Psi_m^* \hat{Q} \Psi_n dV$$

Or, equivalently

$$\int_V (\hat{Q}\Psi_m)^* \Psi_n dV = \int \Psi_m^* \hat{Q} \Psi_n dV$$

Operator that obeys this condition is called a Hermitian operator. In Dirac notation, the integral condition can be simply written as

$$< \hat{Q}\Psi_m|\Psi_n >=< \Psi_m|\hat{Q}\Psi_n >=< \Psi_m|\hat{Q}|\Psi_n > \tag{16}$$

The integral condition can be also rewritten for general functions f and g, where f corresponds to Ψ_m^* and g to Ψ_n.

$$\int_V f\hat{Q}g dV = \int_V g\hat{Q}^* f dV$$

To be complete, we should define what happens when an operator is not Hermitian, i. e. when

$$< \Psi_m|\hat{Q}\Psi_n >\neq< \hat{Q}\Psi_m|\Psi_n >$$

In this case, we define a different operator $\hat{Q}^\dagger$, which we call the Hermitian adjoint, by

$$< \Psi_m|\hat{Q}\Psi_n >=< \hat{Q}^\dagger \Psi_m|\Psi_n > \tag{17}$$

For the special case of Hermitian operators, $\hat{Q}^\dagger = \hat{Q}$, i. e. the Hermitian adjoint of a Hermitian operator is the operator itself. We then say that Hermitian operators are self-adjoint.

Further property can be derived from the definition. Consider case $\Psi_m = \Psi_n$. Then

$$< \Psi_n|\hat{Q}\Psi_n >=< \hat{Q}\Psi_n|\Psi_n >$$

$$\int_V \Psi_n^* \hat{Q}\Psi_n dV = \int_V (\hat{Q}\Psi_n)^* \Psi_n dV$$

Since Ψ_n is an eigenfunction of $\hat{Q}$ (by construction here)

$$\int_V \Psi_n^* q_n \Psi_n dV = \int_V (q_n \Psi_n)^* \Psi_n dV$$

Since q_n is only a constant, it can be taken out of the integrals.

$$q_n \int_V \Psi_n^* \Psi_n dV = q_n^* \int_V \Psi_n^* \Psi_n dV$$

The two sides can only be equal if $q_n^* = q_n$ (for non-zero Ψ_n), which only occurs when q_n is real. This means that Hermitian operators have also real eigenvalues. Furthemore, if we take the general case of two eigenstates of a Hermitian operator

$$< \hat{Q}\Psi_m|\Psi_n >=< \Psi_m|\hat{Q}\Psi_n >$$

$$< q_m \Psi_m|\Psi_n >=< \Psi_m|q_n \Psi_n >$$

$$q_m^* < \Psi_m|\Psi_n > = q_n < \Psi_m|\Psi_n >$$

If the eigenvalues are non-degenerate (different for different m and n), the equation is valid only if $< \Psi_m|\Psi_n > = 0$ for $m \neq n$. This means that non-degenerate eigenstates of Hermitian operators are orthogonal. Now, we show how we can create Hermitian operators from existing Hermitian operators. Trivially, we can show that real (not complex !) constant c is Hermitian

$$\int_V fcgdV = \int_V gcfdV = \int_V gc^*fdV$$

Now consider Hermitian operator $\hat{Q}$. Then operator $\hat{P} = c\hat{Q}$ is also Hermitian, if c is a real (not complex!) constant.

$$\int_V f(c\hat{Q})gdV = c\int_V f\hat{Q}gdV = c\int_V g\hat{Q}^*fdV = \int_V g(c\hat{Q})^*fdV$$

If there are two operators which are Hermitian, $\hat{P}$ and $\hat{Q}$, their sum is also Hermitian

$$\int_V f(\hat{P}+\hat{Q})gdV = \int_V f\hat{P}gdV + \int_V f\hat{Q}gdV = \int_V g\hat{P}^*fdV + \int_V g\hat{Q}^*fdV = \int_V g(\hat{P}^*+\hat{Q}^*)fdV$$

However, the product of two Hermitian operators is not Hermitian

$$\int_V f(\hat{P}\hat{Q})gdV = \int_V f\hat{P}(\hat{Q}g)dV = \int_V (\hat{Q}g)\hat{P}^*fdV = \int_V (\hat{P}^*f)\hat{Q}gdV =$$

$$= \int_V g\hat{Q}^*\hat{P}^*fdV \neq \int_V g(\hat{P}\hat{Q})^*fdV$$

as the change in the order of operators can generaly change the results. However, for two Hermitian operators that can change order, their product is also Hermitian. But this is only very special case of operators, but applies for example for squares of Hermitian operators.

1.6.1   Operator Complex Conjugate

it can be easily proven that complex conjugate of an operator distributes complex conjugation simply algebraically, i. e.

$$((\hat{P}\hat{Q})f)^* = (\hat{P}(\hat{Q}f))^* = \hat{P}^*(\hat{Q}f)^* = \hat{P}^*\hat{Q}^*f^*$$

$$((\hat{P}+\hat{Q})f)^* = (\hat{P}f)^* + (\hat{Q}f)^* = \hat{P}^*f^* + \hat{Q}^*f^* = (\hat{P}^*+\hat{Q}^*)f^*$$

Therefore any operator that is build up from algebraic multiplication and summation (therefore including differential or integral operators) has complex conjugate that is simply algebraic complex conjugation. For example, we search for complex conjugate of $\hat{p}_x = -i\hbar\frac{\partial}{\partial x}$. Lets denote the complex conjugate operator for now $\hat{w}_x$. By our definition

$$(\hat{p}_x f)^* = \hat{w}_x f^*$$

For brevity, consider that f is eigenfunction of $\hat{p}_x$, for example $f = e^{ikx}$ with real k. Then

$$(\hat{p}_x f)^* = \left(-i\hbar\frac{\partial}{\partial x}e^{ikx}\right)^* = \left(-i\hbar(ik)e^{ikx}\right)^* = \hbar k e^{-ikx} = \hat{w}_x e^{-ikx}$$

To satisfy this equation, we can choose $\hat{w}_x = i\hbar\frac{\partial}{\partial x}$, which is simply the algebraic complex conjugate of $\hat{p}_x$. Therefore, finding operator complex conjugates can be done by simply taking their algebraic complex conjugates.

1.7   Commutators

Above, we have used property of operators called associativity - it does not matter how we draw brackets around our operators. Operators are also distributive with respect to addition, which we also used above. However, operators are generally not commuting, which means that switching the order of operators changes the resultant function, i.e.

$$\hat{P}\hat{Q}\Psi \neq \hat{Q}\hat{P}\Psi$$

We then define new operator, that is composite operator of any pair of operators, called the commutator operator, which for operators $\hat{P}$ and $\hat{Q}$ is

$$[\hat{P},\hat{Q}] = \hat{P}\hat{Q} - \hat{Q}\hat{P}$$

We know list a few properties of commutators. First is the anti-symmetry with respect to exchange of operators $\hat{P}$ and $\hat{Q}$, i. e.

$$[\hat{Q}, \hat{P}] = \hat{Q}\hat{P} - \hat{P}\hat{Q} = -(\hat{P}\hat{Q} - \hat{Q}\hat{P}) = -[\hat{P}, \hat{Q}]$$

Other is distribution with respect to addition

$$[\hat{P} + \hat{Q}, \hat{R}] = (\hat{P} + \hat{Q})\hat{R} - \hat{R}(\hat{P} + \hat{Q}) = \hat{P}\hat{R} + \hat{Q}\hat{R} - \hat{R}\hat{P} - \hat{R}\hat{Q} = (\hat{P}\hat{R} - \hat{R}\hat{P}) + (\hat{Q}\hat{R} - \hat{R}\hat{Q}) =$$

$$= [\hat{P}, \hat{R}] + [\hat{Q}, \hat{R}]$$

And the last one is specific identity for product of operators

$$[\hat{P}, \hat{Q}\hat{R}] = \hat{P}\hat{Q}\hat{R} - \hat{Q}\hat{R}\hat{P} = \hat{P}\hat{Q}\hat{R} - \hat{Q}\hat{P}\hat{R} + \hat{Q}\hat{P}\hat{R} - \hat{Q}\hat{R}\hat{P} = (\hat{P}\hat{Q} - \hat{Q}\hat{P})\hat{R} + \hat{Q}(\hat{P}\hat{R} - \hat{R}\hat{P}) =$$

$$= [\hat{P}, \hat{Q}]\hat{R} + \hat{Q}[\hat{P}, \hat{R}]$$

(notice the similarity to the tripple vector expansion, with exception of negative sign)
Now, for the importance of commutators in quantum dynamics. Lets start by considering operators for which

$$[\hat{P}, \hat{Q}] = 0$$

which means that we can exchange the order of operation. Consider an eigenfunction of $\hat{Q}$, Ψ_Q. Then

$$\hat{P}\hat{Q}\Psi_Q = \hat{Q}\hat{P}\Psi_Q$$

$$\hat{P}q\Psi_Q = \hat{Q}\hat{P}\Psi_Q$$

Usually (for multiplicative, differential and integral operators), the constant eigenvalue q is not affected by the operator. Then

$$q\hat{P}\Psi_Q = \hat{Q}\hat{P}\Psi_Q$$

But, this is the eigenvalue equation for operator $\hat{Q}$ operating on function $\hat{P}\Psi_Q$. This means that $\hat{P}\Psi_Q$ is an eigenfunction of $\hat{Q}$. This can only happen if Ψ_Q is an eigenfunction of $\hat{P}$.
We can do similar argument for $\hat{Q}\Psi_P$.

$$\hat{P}\hat{Q}\Psi_P = \hat{Q}\hat{P}\Psi_P = p\hat{Q}\Psi_P$$

hence $\hat{Q}\Psi_P$ is also eigenfunction of $\hat{P}$, which means that Ψ_P is also an eigenfunction of $\hat{Q}$
The commuting operators therefore have a common set of eigenfunctions, and we say that the operators are compatible.

1.8 Uncertainty Principle

Now, we come to a critical topic in quantum mechanics, the uncertainty principle. This principle is fundamental consequence of wave nature of the particles and limits the precision of definition of any two variables. In extreme cases, one variable might be known with perfect precision, while the other has no well defined value.
To discover the reasoning behind uncertainty principle, lets first define the uncertainty in certain observable. We define it as a standard deviation of the value of the observable, which is classically defined as

$$\sigma_Q = \sqrt{<Q^2> - <Q>^2}$$

So

$$\sigma_Q^2 = <Q^2> - <Q>^2$$

Plugging in from our definitions for a general state Ψ

$$<Q^2> = \int_V \Psi^* \hat{Q}^2 \Psi dV$$

$$<Q>^2 = \left(\int_V \Psi^* \hat{Q} \Psi dV \right)^2$$

Now, consider an operator $(\hat{Q}- <Q>)^2$. The expectation of this operator is (because $<Q>$ is just a constant)

$$< (\hat{Q}- <Q>)^2 >= \int_V \Psi^*(\hat{Q}- <Q>)^2 \Psi dV = \int_V \Psi^*(\hat{Q}^2 - 2<Q>\hat{Q}+ <Q>^2)\Psi dV =$$

$$\int_V \Psi^*\hat{Q}^2\Psi dV - 2<Q> \int_V \Psi^*\hat{Q}\Psi dV+ <Q>^2 \int_V \Psi^*\Psi dV$$

For normalized states, this becomes (using definition of expectation value)

$$< (\hat{Q}- <Q>)^2 >=< Q^2 > -2<Q><Q> + <Q>^2=< Q^2 > - <Q>^2= \sigma_Q^2$$

Hence, we found something like an operator of variance/uncertainty - it is operator $(\hat{Q}- <Q>)^2$. I say only something like because operating with this operator does not actually return the variance, rather expectation of this operator is the variance of given variable.

Considering the rules for creation of Hermitian operators, this operator is a square of sum of Hermitian operator $\hat{Q}$ and real constant $- <Q>$, and therefore this operator itself is also Hermitian. This means that we can express it, succinctly in Dirac notation, as

$$< \Psi|(\hat{Q}- <Q>)^2\Psi >=< \Psi|(\hat{Q}- <Q>) \left[(\hat{Q}- <Q>)\Psi \right] >=< (\hat{Q}- <Q>)\Psi|(\hat{Q}- <Q>)\Psi >$$

Since $(\hat{Q}- <Q>)\Psi$ is generally a function, we can write it as function f for brevity, and the variance then is

$$\sigma_Q^2 =< f|f >$$

For some other observable P that is being observed **simultaneously** with variable Q, the variance is

$$\sigma_P^2 =< (\hat{P}- <P>)\Psi|(\hat{P}- <P>)\Psi >$$

Or, in terms of succinct function $g = (\hat{P}- <P>)\Psi$

$$\sigma_P^2 =< g|g >$$

Hence the product of variances of these two variables is

$$\sigma_Q^2\sigma_P^2 =< f|f >< g|g >$$

The last expression is the product of two inner products of elements of Hilbert space with themselves. Using Cauchy-Schwarz inequality (8)

$$\sigma_Q^2\sigma_P^2 =< f|f >< g|g >\geq | < f|g > |^2$$

$$\sigma_Q^2\sigma_P^2 \geq | < (\hat{Q}- <Q>)\Psi|(\hat{P}- <P>)\Psi > |^2 \tag{18}$$

The expression inside the absolute value can be further simplified, since both $\hat{Q}- <Q>$ and $\hat{P}- <P>$ are Hermitian

$$< (\hat{Q}- <Q>)\Psi|(\hat{P}- <P>)\Psi >=< \Psi|(\hat{Q}- <Q>)(\hat{P}- <P>)\Psi >=$$

$$=< \Psi|(\hat{Q}\hat{P}- <P>\hat{Q}- <Q>\hat{P}+ <Q><P>)\Psi >=$$

$$=< \Psi|\hat{Q}\hat{P}\Psi > - <P><\Psi|\hat{Q}|\Psi > - <Q><\Psi|\hat{P}|\Psi > + <Q><P>$$

Using definition of expectation

$$< (\hat{Q}- <Q>)\Psi|(\hat{P}- <P>)\Psi >=< \Psi|\hat{Q}\hat{P}\Psi > - <P><Q> - <Q><P> + <Q><P >=$$

$$=< \Psi|\hat{Q}\hat{P}\Psi > - <P><Q >$$

We would end with the absolute value of this expression, but we can simplify it further if we try to express the absolute value in terms of complex numbers and their conjugates. Generaly, $< f|g >$ is a complex number, for which

$$| < f|g > |^2 = Re(< f|g >)^2 + Im(< f|g >)^2$$

Because both $Re(<f|g>)$ and $Im(<f|g>)$ are real numbers, their squares are positive, and hence

$$|<f|g>|^2 \geq Im(<f|g>)^2$$

Therefore, the original equation becomes

$$\sigma_Q^2 \sigma_P^2 \geq |<f|g>|^2 \geq Im(<f|g>)^2$$

And therefore, taking the positive root, as standard deviations/uncertainties are positive.

$$\sigma_Q \sigma_P \geq |Im(<f|g>)|$$

For any complex number z, we have $Im(z) = \frac{1}{2i}(z - z^*)$, therefore

$$Im(<f|g>) = \frac{1}{2i}(<f|g> - <f|g>^*)$$

The complex conjugate of the inner product is

$$<f|g>^* = \left(\int_V f^* g \, dV \right)^* = \int_V (f^* g)^* dV = \int_V g^* f \, dV = <g|f>$$

Therefore

$$<f|g> - <f|g>^* = <f|g> - <g|f> =$$
$$= <(\hat{Q} - <Q>)\Psi|(\hat{P} - <P>\Psi) - <(\hat{P} - <P>)\Psi|(\hat{Q} - <Q>)\Psi> =$$
$$<\Psi|(\hat{Q} - <Q>)(\hat{P} - <P>)\Psi> - <\Psi|(\hat{P} - <P>)(\hat{Q} - <Q>)\Psi> =$$
$$= <\Psi|(\hat{Q}\hat{P} - <Q>\hat{P} - <P>\hat{Q} + <P><Q>)\Psi>$$
$$- <\Psi|(\hat{P}\hat{Q} - <Q>\hat{P} - <P>\hat{Q} + <P><Q>)\Psi> =$$
$$= <\Psi|\hat{Q}\hat{P}\Psi> - <\Psi|\hat{P}\hat{Q}\Psi> = <\Psi|(\hat{Q}\hat{P} - \hat{P}\hat{Q})\Psi> = <\Psi|[\hat{Q}, \hat{P}]\Psi>$$

We can show that the commutator of two Hermitian operators is in some sense anti-Hermitian, i. e.

$$\int_V \Psi^*[\hat{Q}, \hat{P}]\Psi dV = \int_V \Psi^*(\hat{Q}\hat{P} - \hat{P}\hat{Q})\Psi dV = \int_V \Psi^*\hat{Q}\hat{P}\Psi dV - \int_V \Psi^*\hat{P}\hat{Q}\Psi dV =$$

$$= \int_V \hat{Q}^*\Psi^*\hat{P}\Psi dV - \int_V \hat{P}^*\Psi^*\hat{Q}\Psi dV = \int_V \hat{P}^*\hat{Q}^*\Psi^*\Psi dV - \int_V \hat{Q}^*\hat{P}^*\Psi^*\Psi dV =$$

$$= \int_V (\hat{P}^*\hat{Q}^* - \hat{Q}^*\hat{P}^*)\Psi^*\Psi dV = -\int_V [\hat{Q}, \hat{P}]^*\Psi^*\Psi dV$$

So, in Dirac notation, we have

$$<\Psi|[\hat{Q}, \hat{P}]\Psi> = -<[\hat{Q}, \hat{P}]\Psi|\Psi>$$

But, in absolute values

$$|<\Psi|[\hat{Q}, \hat{P}]\Psi>| = |<[\hat{Q}, \hat{P}]\Psi|\Psi>|$$

Therefore, we can label this state as $|<[\hat{Q}, \hat{P}]>|$, in analogy with expectation values. Furthemore, since $[\hat{Q}, \hat{P}] = -[\hat{P}, \hat{Q}]$, also

$$<[\hat{Q}, \hat{P}]> = |<\Psi|[\hat{Q}, \hat{P}]\Psi>| = |<\Psi|[\hat{P}, \hat{Q}]\Psi>| = <[\hat{P}, \hat{Q}]>$$

Therefore, plugging all together (since absolute value distributes for products)

$$|Im(<f|g>)| = \left| \frac{1}{2i}(<f|g> - <g|f>) \right| = \left| \frac{1}{2i} \right| |<f|g> - <g|f>| = \left| \frac{1}{2i} <[\hat{Q}, \hat{P}]> \right|$$

And thus

$$\sigma_Q \sigma_P \geq \left| \frac{1}{2i} <[\hat{Q}, \hat{P}]> \right| \tag{19}$$

which is a general uncertainty principle.

In an example, consider a momentum operator $\hat{p}_x$ and position operator $\hat{x}$. The commutator is (with explicitly written function)

$$[\hat{p}_x, \hat{x}]f = -i\hbar \frac{\partial}{\partial x}xf - x\left(-i\hbar \frac{\partial}{\partial x} \right)f = -i\hbar f - i\hbar x \frac{\partial f}{\partial x} + i\hbar x \frac{\partial f}{\partial x} = -i\hbar f$$

$$[\hat{p}_x, \hat{x}] = -i\hbar$$

Hence

$$\sigma_{p_x} \sigma_x \geq \left| \frac{1}{2i}(-i\hbar) \right| = \frac{\hbar}{2} \tag{20}$$

which is classical Heisenberg uncertainty principle.

1.9   Time Dependence of Expectation Values

Last topic in very general QM discovered here is the time dependence of expectation values for stationary states (states that are solutions to TISE). The time dependence of the expectation, if there is any, depends on the derivative with respect to time. From definition of expectation value

$$\frac{d<Q>}{dt} = \frac{d}{dt}\int_V \Psi^* \hat{Q} \Psi dV = \int_V \frac{d\Psi^*}{dt}\hat{Q}\Psi dV + \int_V \Psi^* \frac{d\hat{Q}\Psi}{dt}dV$$

The only time dependence that can be added by the operator to Ψ is if the operator itself is explicitly time dependent (function of t). If the operator is time dependent only explicitly, the total derivative is equal to the partial derivative, so

$$\frac{d\hat{Q}}{dt} = \frac{\partial \hat{Q}}{\partial t}$$

Since Ψ is stationary state solution, $\Psi = \psi(\vec{r})\rho(t)$ and hence

$$\frac{d\Psi}{dt} = \psi(\vec{r})\frac{d\rho}{dt} = \psi(\vec{r})\frac{\partial \rho}{\partial t} = \frac{\partial \Psi}{\partial t}$$

Similarly for the complex conjugate,

$$\frac{d\Psi}{dt} = \frac{\partial \Psi}{\partial t}$$

Hence the time dependence of expectation value becomes

$$\frac{d<Q>}{dt} = \int_V \frac{\partial \Psi^*}{\partial t}\hat{Q}\Psi dV + \int_V \Psi^* \frac{\partial \hat{Q}}{\partial t}\Psi dV + \int_V \Psi^* \hat{Q}\frac{\partial \Psi}{\partial t}dV$$

TDSE is

$$\hat{H}\Psi = i\hbar \frac{\partial \Psi}{\partial t}$$

$$\frac{\partial \Psi}{\partial t} = -\frac{i}{\hbar}\hat{H}\Psi$$

And

$$\frac{\partial \Psi^*}{\partial t} = \frac{i}{\hbar}\hat{H}^*\Psi^*$$

Hence we have

$$\frac{d<Q>}{dt} = \int_V \frac{i}{\hbar}\hat{H}^*\Psi^*\hat{Q}\Psi dV + \int_V \Psi^* \frac{\partial \hat{Q}}{\partial t}\Psi dV + \int_V \Psi^* \hat{Q}\left(-\frac{i}{\hbar}\hat{H}\Psi\right)dV =$$

$$\left\langle \frac{\partial \hat{Q}}{\partial t}\right\rangle + \frac{i}{\hbar}\left(\int_V \hat{H}^*\Psi^*\hat{Q}\Psi dV - \int_V \Psi^*\hat{Q}\hat{H}\Psi dV\right)$$

Since $\hat{H}$ is hermitian (total energy is always real)

$$\frac{d<Q>}{dt} = \left\langle \frac{\partial \hat{Q}}{\partial t}\right\rangle + \frac{i}{\hbar}\left(\int_V \Psi^*\hat{H}\hat{Q}\Psi dV - \int_V \Psi^*\hat{Q}\hat{H}\Psi dV\right) = \left\langle \frac{\partial \hat{Q}}{\partial t}\right\rangle + \frac{i}{\hbar} < \Psi|[\hat{H},\hat{Q}]\Psi >$$

Therefore, if there is no explicit time dependence in the operator of an observable, the expectation value of the observable only changes in time if the observable does not commute with the Hamiltonian operator of the system, and the change is proportional to the expectation of the commutator of $\hat{H}$ and $\hat{Q}$.

2   Formal Applications of Quantum Mechanics

In this section, some more formal applications of quantum mechanical formalism will be presented. These try to stick to principles of quantum mechanics as closely as possible. First is a discussion of a general system that can exist in only two observable states.

2.1 Two State System

Let Ψ_1 and Ψ_2 be the observable states of the system. These are linear combinations of the eigenstates of the hamiltonian operator $\hat{H}$. Let Φ_1 and Φ_2 be these energy eigenstates. There are only two since we have to be able to also express these in terms of linear combination of the observable states, which there is only two of and therefore can only form two other linearly independent states. Now, lets represent the states as vectors with components of vectors equal to the coefficient in linear superposition of observable eigenstates that creates the state, i.e. for general state Ψ

$$\Psi = c_1\Psi_1 + c_2\Psi_2 = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

The general states than obeys Schrödinger equation as discussed before

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} <\Psi_1|\hat{H}|\Psi_1> & <\Psi_1|\hat{H}|\Psi_2> \\ <\Psi_2|\hat{H}|\Psi_1> & <\Psi_2|\hat{H}|\Psi_2> \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \mathbf{H}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

Let

$$\Phi_1 = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$$

where p_1 and p_2 are such that Φ_1 is always eigenvector of $\mathbf{H}$, which usually means these are constant in time (if Hamiltonian is not explicitly dependent on time).
Analogously for Φ_2

$$\Phi_2 = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}$$

For non-degenerate matrix $\mathbf{H}$, we can express the general state vector in terms of the eigenvectors as

$$\Psi = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = k_1\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} + k_2\begin{pmatrix} q_1 \\ q_2 \end{pmatrix}$$

where we allow for time dependence of the k coefficients, which enables the state vector to change. The Schrödinger equation then is

$$i\hbar\frac{\partial}{\partial t}\vec{c} = \mathbf{H}(k_1\Phi_1 + k_2\Phi_2)$$

$$i\hbar\frac{\partial}{\partial t}(k_1\Phi_1 + k_2\Phi_2) = k_1\phi_1\Phi_1 + k_2\phi_2\Phi_2$$

where $\phi_{1/2}$ are the eigenvalues of $\Phi_{1/2}$.
This is another effectively vector equation (for constant energy eigenstates), but we need to change the basis to the basis of eigenvectors of energy Φ_1 and Φ_2. This equation than is

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix} k_1(t) \\ k_2(t) \end{pmatrix} = \begin{pmatrix} \phi_1 k_1(t) \\ \phi_2 k_2(t) \end{pmatrix}$$

Since ks only depend on time, the equation is equivalent with

$$\begin{pmatrix} \frac{dk_1}{dt} \\ \frac{dk_2}{dt} \end{pmatrix} = \begin{pmatrix} \frac{-i}{\hbar}\phi_1 k_1 \\ \frac{-i}{\hbar}\phi_2 k_2 \end{pmatrix}$$

This has solution

$$\begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = \begin{pmatrix} k_{10}e^{-i\frac{\phi_1}{\hbar}t} \\ k_{20}e^{-i\frac{\phi_2}{\hbar}t} \end{pmatrix}$$

where k_{10} and k_{20} are some constant amplitudes. This means that for real eigenvalues of energy, each components rotates in complex plane, each with different velocity. This than causes the general state to change as (in the original basis of observable states Ψ_1 and Ψ_2)

$$\Psi = k_{10}e^{-i\frac{\phi_1}{\hbar}t}(p_1\Psi_1 + p_2\Psi_2) + k_{20}e^{-i\frac{\phi_2}{\hbar t}}(q_1\Psi_1 + q_2\Psi_2) =$$

$$= \left(k_{10}e^{-i\frac{\phi_1}{\hbar}t}p_1 + k_{20}e^{-i\frac{\phi_2}{\hbar}t}q_1\right)\Psi_1 + \left(k_{10}e^{-i\frac{\phi_1}{\hbar}t}p_2 + k_{20}e^{-i\frac{\phi_2}{\hbar}t}q_2\right)\Psi_2 = a\Psi_1 + b\Psi_2$$

In order for this general state to be normalized, assuming that the observable states are orthonormal (and as they have to be orthogonal, this it is effectively only requirement for them to be normalized)

$$< \Psi | \Psi >= 1$$

$$< a\Psi_1 + b\Psi_2 | a\Psi_1 + b\Psi_2 >= a^*a < \Psi_1 | \Psi_1 > + a^*b < \Psi_1 | \Psi_2 > + ab^* < \Psi_2 | \Psi_1 > + bb^* < \Psi_2 | \Psi_2 >= 1$$

$$|a|^2 + |b|^2 = 1$$

$$|k_{10}|^2(|p_1|^2 + |p_2|^2) + k_{10}k_{20}^*(p_1q_1^* + p_2q_2^*)e^{i\frac{\phi_2 - \phi_1}{\hbar}t} + k_{10}^*k_{20}(p_1^*q_1 + p_2^*q_2)e^{i\frac{\phi_1 - \phi_2}{\hbar}t} + |k_{20}|^2(|q_1|^2 + |q_2|^2) = 1$$

This is quite a complicated expression. It can be slightly simplified by noticing that the inner two expressions are complex conjugates, and therefore it becomes

$$|k_{10}|^2(|p_1|^2 + |p_2|^2) + |k_{20}|^2(|q_1|^2 + |q_2|^2) + 2Re\left(k_{10}k_{20}^*(p_1q_1^* + p_2q_2^*)e^{i\frac{\phi_2 - \phi_1}{\hbar}t}\right) = 1$$

Often, we are interested in a case when the state is at some time $t = 0$ prepared in one of the observable states. For example, say that the state is prepared in a state Ψ_1. Then, we have

$$\Psi(t = 0) = (k_{10}p_1 + k_{20}q_1)\Psi_1 + (k_{10}p_2 + k_{20}q_2)\Psi_2 = \Psi_1$$

Which requires

$$k_{10}p_1 = 1 - k_{20}q_1$$

and

$$k_{10}p_2 = -k_{20}q_2$$

This is a set of simultaneous equations. From the second one

$$k_{10} = -\frac{q_2}{p_2}k_{20}$$

Hence, by substituting into the first one

$$-\frac{q_2}{p_2}p_1k_{20} = 1 - q_1k_{20}$$

$$\left(q_1 - \frac{p_1}{p_2}q_2\right)k_{20} = 1$$

$$k_{20} = \frac{p_2}{q_1p_2 - p_1q_2}$$

And consequently

$$k_{10} = \frac{-q_2}{q_1p_2 - p_1q_2}$$

These equations can be then used to express the state using just eigenvalues of energies and the components of the eigenvectors. Now, a few examples of this formalism follow, but they are preceded by mathematical intermezzo on finding eigenvalues and eigenvectors.

2.1.1   Matrix Eigenvalues and Eigenvectors

Let $\mathbf{A}$ be a matrix and $\vec{c}$ an eigenvector of the matrix. Then

$$\mathbf{A}\vec{c} = \lambda\vec{c}$$

where scalar λ is the eigenvalue corresponding to $\vec{c}$. This means that

$$\mathbf{A}\vec{c} - \lambda c = \vec{0}$$

where $\vec{0}$ is the zero vector.

$$\mathbf{A}\vec{c} - \lambda\mathbf{I}\vec{c} = \vec{0}$$

where $\mathbf{I}$ is the identity matrix. Then

$$(\mathbf{A} - \lambda\mathbf{I})\vec{c} = \vec{0}$$

If we do not want $\vec{c}$ trivial ($\vec{0}$), then we must have determinant of the matrix $\mathbf{A} - \lambda\mathbf{I}$ equal to zero

$$|\mathbf{A} - \lambda\mathbf{I}| = 0$$

This is the main equation with which to search for eigenvectors. However, there exists a few tricks how to make searching slightly easier.

(Adding identity matrix)

Assume we know some eigenvector $\vec{c}$ of matrix $\mathbf{A}$. We now have a different matrix $\mathbf{B}$, which is $\mathbf{B} = \mathbf{A} + b\mathbf{I}$, where b is some constant, and we want to know its eigenvalues and eigenvectors, $\vec{d}$. For these

$$\mathbf{B}\vec{d} = \phi\vec{d}$$

where ϕ is the eigenvalue of $\mathbf{B}$. This means

$$(\mathbf{B} - \phi\mathbf{I})\vec{d} = 0$$

and, substituting from definition of $\mathbf{B}$

$$(\mathbf{A} + b\mathbf{I} - \phi\mathbf{I})\vec{d} = 0$$

$$(\mathbf{A} - (\phi - b)\mathbf{I})\vec{d} = 0$$

But, this is an equation for eigenvalue of $\mathbf{A}$, so the eigenvectors of $\mathbf{B}$ are the same as eigenvectors of $\mathbf{A}$, and the eigenvalues are

$$\lambda = \phi - b$$

$$\phi = \lambda + b$$

So, the eigenvalues of $\mathbf{B}$ are increased by b.

2.1.2   Amonia Molecule

The amonia molecule can exist in two energically equivalent states, however, switching between these states can occur if the states overcome some potential barrier.

We state that the two states of amonia molecule are not the eigenstates of the hamiltonian operator, and we state that the hamiltonian matrix is

$$\mathbf{H} = \begin{pmatrix} \epsilon_0 & V \\ V & \epsilon_0 \end{pmatrix}$$

The eigenvectors will be the same as for matrix

$$\mathbf{H}' = \begin{pmatrix} 0 & V \\ V & 0 \end{pmatrix}$$

Finding eigenvalues of $\mathbf{H}'$ (not of $\mathbf{H}$)

$$\begin{vmatrix} -\lambda' & V \\ V & -\lambda' \end{vmatrix} = 0$$

$$(\lambda')^2 - V^2 = 0$$

$$\lambda' = \pm V$$

Which means that eigenvalues of $\mathbf{H}$ are

$$\lambda = \epsilon_0 \pm V$$

Finding the eigenvectors of $\mathbf{H}'$ (which are the same as of $\mathbf{H}$)

$$\begin{pmatrix} \mp V & V \\ V & \mp V \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

This is a system of two simultaneous equations

$$bV \mp aV = 0$$

$$aV \mp bV = 0$$

From the first equation

$$b = \pm a$$

So the two eigenvectors are

$$\vec{p} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \vec{q} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

with corresponding eigenvalues

$$\phi_1 = \epsilon_0 + V, \phi_2 = \epsilon_0 - V$$

We will now search for time in which the molecule transitions from one measurable state to the other state. At time $t = 0$, the molecule starts in state 1.

Looking now back to our previous formalism, the mixing coefficients k_{10} and k_{20} are

$$k_{10} = \frac{-q_2}{q_1 p_2 - p_1 q_2} = \frac{-(-1)}{1 - (-1)} = \frac{1}{2}$$

$$k_{20} = \frac{p_2}{q_1 p_2 - p_1 q_2} = \frac{1}{2}$$

So, the general state of the system is given as

$$\Psi = k_{10} e^{-i\frac{\phi_1}{\hbar}t}\vec{p} + k_{20}e^{-i\frac{\phi_2}{\hbar}t}\vec{q} = \frac{1}{2}\left(e^{-i\frac{\phi_1}{\hbar}t}\begin{pmatrix}1\\1\end{pmatrix} + e^{-i\frac{\phi_2}{\hbar}t}\begin{pmatrix}1\\-1\end{pmatrix}\right) =$$

$$= \frac{1}{2}\left(\left(e^{-i\frac{(\epsilon_0+V)}{\hbar}t} + e^{-i\frac{(\epsilon_0-V)}{\hbar}t}\right)\begin{pmatrix}1\\0\end{pmatrix} + \left(e^{-i\frac{(\epsilon_0+V)}{\hbar}t} - e^{-i\frac{(\epsilon_0-V)}{\hbar}t}\right)\begin{pmatrix}0\\1\end{pmatrix}\right) =$$

$$= \frac{1}{2}\left(e^{-i\frac{\epsilon_0}{\hbar}t}\left(e^{-i\frac{V}{\hbar}t} + e^{i\frac{V}{\hbar}t}\right)\Psi_1 + e^{-i\frac{\epsilon_0}{\hbar}t}\left(e^{-i\frac{V}{\hbar}t} - e^{i\frac{V}{\hbar}t}\right)\Psi_2\right) =$$

$$= e^{-i\frac{\epsilon_0}{\hbar}t}\left(\cos\left(\frac{V}{\hbar}t\right)\Psi_1 - i\sin\left(\frac{V}{\hbar}t\right)\Psi_2\right)$$

This means that at time $t = \frac{\hbar}{V}\frac{\pi}{2}$, the general state does not include any part of the initial state Ψ_1, and t is the transition time.

2.2   Fermions and Bosons and Exchange symmetry

For systems with more than one particle, we can still assign one wavefunction to the whole system, which than represents the properties of the whole system (in a similar way as we can write a hamiltonian for a whole system).

Consider that we now have two identical particles. The wavefunction than depends on position of these two particles, and time. So

$$\Psi = \Psi(\vec{r}_1, \vec{r}_2, t)$$

Lets define exchange operator $\hat{P}$ which switches the first particle with the second particle, i. e. does transformation

$$\hat{P}\Psi(\vec{r}_1, \vec{r}_2, t) = \Psi(\vec{r}_2, \vec{r}_1, t)$$

Because the particles are indistinguishable, the probability density profile must not change after the operation. This means that

$$|\Psi(\vec{r}_1, \vec{r}_2, t)|^2 = |\Psi(\vec{r}_2, \vec{r}_1, t)|^2$$

This leaves us with two possibilities. Either

$$\Psi(\vec{r}_1, \vec{r}_2, t) = -\Psi(\vec{r}_2, \vec{r}_1, t)$$

Or

$$\Psi(\vec{r}_1, \vec{r}_2, t) = \Psi(\vec{r}_2, \vec{r}_1, t)$$

This effectively means that $\hat{P}$ has eigenvalues ± 1. Now, consider that two particles are exchange antisymmetric and have the same position $\vec{r}$. Then

$$\hat{P}\Psi(\vec{r}_1 = \vec{r}, \vec{r}_2 = \vec{r}, t) = \Psi(\vec{r}_2 = \vec{r}, \vec{r}_1 = \vec{r}, t)$$

$$-\Psi(\vec{r}, \vec{r}, t) = \Psi(\vec{r}, \vec{r}, t)$$

This means that $\Psi = 0$, or that the probability of observing the two particles at the same position is zero. This can be further extanded to any variable that defines the state of the particle, which means that particles which are exchange anti-symmetric cannot exist in the same state simultaneously. For example, if the state of the particle is defined by four quantum numbers as $|n, l, m_l, s>$, then

$$\Psi = |n_1, l_1, m_{l1}, s_1, n_2, l_2, m_{l2}, s_2 >$$

$$\hat{P}_n\Psi = |n_2, l_1, m_{l1}, s_1, n_1, l_2, m_{l2}, s_2 >= -\Psi = -|n_1, l_1, m_{l1}, s_1, n_2, l_2, m_{l2}, s_2 >$$

For $n_1 = n_2$, we again have $\Psi = 0$, and we could define respective operators for each of quantum numbers. Such particles that cannot exist in common state are called fermions.

For the exchange symmetric particles, there is no such obstruction. These particles are called bosons.

2.3   Harmonic Oscillator

Harmonic oscillator is a system of particles that has a parabolic potential. It is generally important because any local minimum in any complicated potential landscape can be modelled for small excitations as parabolic potential.

This can be easily seen by applying Taylor series in 1D. Let $V(x)$ be the potential function, and let x_0 be the position of a local minimum in V and dx small step from the minimum position. The value of the potential at $x_0 + dx$ is

$$V(x_0 + dx) = V(x_0) + \frac{dV}{dx}\big|_{x_0}(x - x_0) + \frac{1}{2}\frac{d^2V}{dx^2}\big|_{x_0}(x - x_0)^2$$

Since the position is minimum, $\frac{dV}{dx}\big|_{x_0} = 0$. By also changing the origin so that x_0 is at the origin, we have

$$V(dx) = V(0) + \frac{1}{2}\frac{d^2V}{dx^2}\big|_0 x^2$$

which is a parabolic potential. Usually, we mark constant derivative term as $k = \frac{d^2V}{dx^2}$ and choose the potential level such that $V(0) = 0$.

The Hamiltonian operator for harmonic oscillator is then (in 1D)

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2}kx^2$$

2.3.1   Ladder Operators

We can try to factorize the hamiltonian operator as

$$\frac{1}{2m}\left(\sqrt{km}x + \hbar\frac{\partial}{\partial x}\right)\left(\sqrt{km}x - \hbar\frac{\partial}{\partial x}\right) = \frac{1}{2m}\left(kmx^2 - \sqrt{km}x\frac{\partial}{\partial x} + \hbar\frac{\partial}{\partial x}(\sqrt{km}x) - \hbar^2\frac{\partial^2}{\partial x^2}\right) =$$

$$= \frac{1}{2m}\left(kmx^2 + \hbar\sqrt{km} - \hbar^2\frac{\partial^2}{\partial x^2}\right) = \frac{1}{2}kx^2 - \hbar^2\frac{\partial^2}{\partial x^2} + \frac{1}{2}\hbar\sqrt{\frac{k}{m}} = \hat{H} + \frac{1}{2}\hbar\omega_0$$

where $\omega_0 = \sqrt{\frac{k}{m}}$ is the classical natural frequency of the harmonic oscillator. Similarly, we could factorize as

$$\frac{1}{2m}\left(\sqrt{km}x - \hbar\frac{\partial}{\partial x}\right)\left(\sqrt{km}x + \hbar\frac{\partial}{\partial x}\right) = \frac{1}{2m}\left(kmx^2 - \hbar^2\frac{\partial^2}{\partial x^2} + \sqrt{km}x\hbar\frac{\partial}{\partial x} - \hbar\frac{\partial}{\partial x}\sqrt{km}x\right) = \hat{H} - \frac{1}{2}\hbar\omega_0$$

Labeling (for reasons that will become clear later)

$$a_+ = \frac{1}{\sqrt{2m}}\left(\sqrt{km}x - \hbar\frac{\partial}{\partial x}\right) \tag{21}$$

$$a_- = \frac{1}{\sqrt{2m}}\left(\sqrt{km}x + \hbar\frac{\partial}{\partial x}\right) \tag{22}$$

We can rewrite the hamiltonian operator as

$$\hat{H} = a_-a_+ - \frac{1}{2}\hbar\omega_0 = a_+a_- + \frac{1}{2}\hbar\omega_0 \tag{23}$$

Besides other applications, we can directly determine the commutator $[a_-, a_+]$

$$[a_-, a_+] = a_-a_+ - a_+a_- = \hbar\omega_0$$

Now, consider that we have some eigenstate of hamiltonian operator Ψ_n. For stationary state, TISE is

$$\hat{H}\Psi_n = E_n\Psi_n$$

$$\left(a_-a_+ - \frac{1}{2}\hbar\omega_0\right)\Psi_n = E_n\Psi_n$$

Operating with a_+ on both sides of the equation

$$\left(a_+ a_- a_+ - \frac{1}{2}\hbar\omega_0 a_+\right)\Psi_n = E_n a_+ \Psi_n$$

Labeling $\Psi_{n+1} = a_+ \Psi_n$, we have

$$\left(a_+ a_- - \frac{1}{2}\hbar\omega_0\right)\Psi_{n+1} = E_n \Psi_{n+1}$$

$$\left(a_+ a_- + \hbar\omega_0 - \frac{1}{2}\hbar\omega\right)\Psi_{n+1} = (E_n + \hbar\omega_0)\Psi_{n+1}$$

$$\left(a_+ a_- + \frac{1}{2}\hbar\omega_0\right)\Psi_{n+1} = (E_n + \hbar\omega_0)\Psi_{n+1}$$

$$\hat{H}\Psi_{n+1} = (E_n + \hbar\omega_0)\Psi_{n+1}$$

Hence, Ψ_{n+1} is also an eigenfunction of the hamiltonian operator, with eigenvalue increased by $\hbar\omega_0$. Similarly, we could start from

$$\hat{H}\Psi_n = E_n \Psi_n$$

$$\left(a_+ a_- + \frac{1}{2}\hbar\omega_0\right)\Psi_n = E_n \Psi_n$$

Now apply a_- to both sides

$$\left(a_- a_+ a_- + \frac{1}{2}\hbar\omega_0 a_-\right)\Psi_n = E_n a_- \Psi_n$$

Setting $\Psi_{n-1} = a_- \Psi_n$

$$\left(a_- a_+ + \frac{1}{2}\hbar\omega_0 - \hbar\omega\right)\Psi_{n-1} = (E_n - \hbar\omega_0)\Psi_{n-1}$$

$$\hat{H}\Psi_{n-1} = (E_n - \hbar\omega_0)\Psi_{n-1}$$

Which means that Ψ_{n-1} is also eigenfunction of the hamiltonian operator.
But, there is a minimal energy eigenvalue that physically makes sense, that is for kinetic energy equal to zero and potential energy equal to minimum $V(0) = 0$. So the total energy must be always greater then zero. This means that when applying a_- to some function Ψ_0 which has energy eigenvalue smaller than $\hbar\omega_0$, we should not get a physical result, as well as for any other consecutive application of a_-. A way to ensure this is to set

$$a_- \Psi_0 = 0 \tag{24}$$

$$\frac{1}{\sqrt{2m}}\left(\sqrt{km}x + \hbar\frac{\partial}{\partial x}\right)\Psi_0 = 0$$

For Ψ_0 only in 1D, this becomes differential equation

$$\frac{d\Psi_0}{dx} = -\frac{\sqrt{km}}{\hbar}x\Psi_0$$

Which has solution

$$\Psi_0 = Ae^{-\frac{\sqrt{km}}{2\hbar}x^2} = Ae^{-\frac{m\omega_0}{2\hbar}x^2}$$

where A is some normalization constant.
Therefore, we have effectively found all eigenfunctions of the hamiltonian operator. Other functions can be simply found by operating with a_+ operator on the Ψ_0 function. To find all the eigenvalues, we need to find the energy eigenvalue E_0 of the eigenfunction Ψ_0. This can be also found from our "boundary condition"

$$a_- \Psi_0 = 0$$

$$a_+ a_- \Psi_0 = 0$$

$$\left(a_+ a_- + \frac{1}{2}\hbar\omega_0\right)\Psi_0 = \frac{1}{2}\hbar\omega_0 \Psi_0$$

$$\hat{H}\Psi_0 = \frac{1}{2}\hbar\omega_0 \Psi_0 = E_0 \Psi_0$$

which means that $E_0 = \frac{1}{2}\hbar\omega_0$ and $E_n = n\hbar\omega + \frac{1}{2}\hbar\omega = \hbar\omega\left(n + \frac{1}{2}\right)$.

The general formula for the eigenfunction can be derived as

$$\Psi_n = C_n H_n \left(\sqrt{\frac{m\omega_0}{\hbar}} x\right) e^{-\frac{m\omega_0}{2\hbar} x^2}$$

where C_n is the normalisation constant and $H_n(x)$ are the Hermite polynomials.

The important message to take is that the energy of the Harmonic oscillator is never truly zero - the minimum value is $\frac{1}{2}\hbar\omega_0$. Also, since Gaussian function $e^{-\alpha x^2}$ has finite value for any point x, the particle has non-zero probability of existing in places where it would be otherwise classicaly forbidden, i. e. at places where the potential energy of the particle is greater than its total energy in the minimum.

2.3.2 Molecular Oscillations

We can model the molecular vibrations as harmonic oscillators with reduced mass μ instead of mass m, for which

$$\mu = \frac{m_1 m_2}{m_1 + m_2}$$

we can see that this mass reduces to m_2 for $m_1 >> m_2$, as in such case only the second molecule oscillates (the second does not move).

The "spring constant" k for such oscillators is usually very high (about 10^3 N m^{-1}). When photons of frequency that is equal to an integer multiple of natural frequency of the oscillator are incident on these oscillators/molecules, they can get absorbed and the energy of the oscillator increases by given $n\hbar\omega_0$. This can be used in some spectroscopy experiments.

2.4 Angular Momentum Operators

We defined angular momentum operators in cartesian coordinates as

$$\hat{L}_x = y\hat{p}_z - z\hat{p}_y$$

$$\hat{L}_y = z\hat{p}_x - x\hat{p}_z$$

$$\hat{L}_z = x\hat{p}_y - y\hat{p}_x$$

We now further define a angular momentum magnitude operator, which we define, by principle of correspondence, as

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$$

where $\hat{L}^2$ should be really $\widehat{(L^2)}$, as L^2 is the name of the observable we are relating to this operator.

2.4.1 Self Commutation

We will now be interested whether we can now all components of angular momentum at the same time. We know that for linear momentum operators, this is possible as

$$[\hat{p}_x, \hat{p}_y] = [\hat{p}_x, \hat{p}_z] = [\hat{p}_y, \hat{p}_z] = 0$$

We also know that

$$[x, \hat{p}_x] = x(-i\hbar)\frac{\partial}{\partial x} - (-i\hbar)\frac{\partial}{\partial x}x = (-i\hbar)\left(x\frac{\partial}{\partial x} - 1 - x\frac{\partial}{\partial x}\right) = i\hbar$$

and similarly $[y, \hat{p}_y] = i\hbar$ and $[z, \hat{p}_z] = i\hbar$.

We now calculate the commutators of the angular momentum components. Using the relations stated before

$$[\hat{L}_x, \hat{L}_y] = [y\hat{p}_z - z\hat{p}_y, z\hat{p}_x - x\hat{p}_z] = [y\hat{p}_z, z\hat{p}_x - x\hat{p}_z] - [z\hat{p}_y, z\hat{p}_x - x\hat{p}_z] =$$

$$= [z\hat{p}_x - x\hat{p}_z, z\hat{p}_y] - [z\hat{p}_x - x\hat{p}_z, y\hat{p}_z] = [z\hat{p}_x, z\hat{p}_y] - [x\hat{p}_z, z\hat{p}_y] - [z\hat{p}_x, y\hat{p}_z] + [x\hat{p}_z, y\hat{p}_z]$$

Taking these components separately

$$[z\hat{p}_x, z\hat{p}_y] = [z\hat{p}_x, z]\hat{p}_y + z[z\hat{p}_x, \hat{p}_y] = -[z, z\hat{p}_x]\hat{p}_y - z[\hat{p}_y, z\hat{p}_x] =$$

$$= -[z, z]\hat{p}_x\hat{p}_y - z[z, \hat{p}_x]\hat{p}_y - z[\hat{p}_y, z]\hat{p}_x - z^2[\hat{p}_y, \hat{p}_x]$$

Now, let $x = x_1, y = x_2, z = x_3$. For any two components $x_i, x_j; i, j \in \{1, 2, 3\}$ $[x_i, x_j] = 0$ since this is only a multiplication, which always commutes. Also, for any $i \neq j; i, j \in \{1, 2, 3\}$

$$[x_i, p_{x_j}] = x_i(-i\hbar)\frac{\partial}{\partial x_j} - (-i\hbar)\frac{\partial}{\partial x_j}x_i = 0$$

as x_i is effectively a constant to differentiation with respect to x_j.

Therefore

$$[z\hat{p}_x, z\hat{p}_y] = 0$$

Second term is

$$-[x\hat{p}_z, z\hat{p}_y] = -[x\hat{p}_z, z]\hat{p}_y - z[x\hat{p}_z, \hat{p}_y] = [z, x\hat{p}_z]\hat{p}_y + z[\hat{p}_y, x\hat{p}_z] =$$
$$= [z, x]\hat{p}_z\hat{p}_y + x[z, \hat{p}_z]\hat{p}_y + zx[\hat{p}_y, \hat{p}_z] + z[\hat{p}_y, x]\hat{p}_z = i\hbar x\hat{p}_y$$

Third term is

$$-[z\hat{p}_x, y\hat{p}_z] = -[z\hat{p}_x, y]\hat{p}_z - y[z\hat{p}_x, \hat{p}_z] = [y, z\hat{p}_x]\hat{p}_z + y[\hat{p}_z, z\hat{p}_x] =$$
$$= [y, z]\hat{p}_x\hat{p}_z + z[y, \hat{p}_x]\hat{p}_z + y[\hat{p}_z, z]\hat{p}_x + yz[\hat{p}_z, \hat{p}_x] = -y[z, \hat{p}_z]\hat{p}_x = -i\hbar y\hat{p}_x$$

And the last term is

$$[x\hat{p}_z, y\hat{p}_z] = [x\hat{p}_z, y]\hat{p}_z + y[x\hat{p}_z, \hat{p}_z] = -[y, x\hat{p}_z]\hat{p}_z - y[\hat{p}_z, x\hat{p}_z] =$$
$$= -[y, x]\hat{p}_z\hat{p}_z - x[y, \hat{p}_z]\hat{p}_z - y[\hat{p}_z, x]\hat{p}_z - yx[\hat{p}_z, \hat{p}_z] = 0$$

And therefore

$$[\hat{L}_x, \hat{L}_y] = 0 + i\hbar x\hat{p}_y - i\hbar y\hat{p}_x + 0 = i\hbar\hat{L}_z$$

Similarly for other commutators

$$[\hat{L}_y, \hat{L}_z] = [z\hat{p}_x - x\hat{p}_z, x\hat{p}_y - y\hat{p}_x] = [z\hat{p}_x, x\hat{p}_y - y\hat{p}_x] - [x\hat{p}_z, x\hat{p}_y - y\hat{p}_x] =$$
$$= [x\hat{p}_y - y\hat{p}_x, x\hat{p}_z] - [x\hat{p}_y - y\hat{p}_x, z\hat{p}_x] = [x\hat{p}_y, x\hat{p}_z] - [y\hat{p}_x, x\hat{p}_z] - [x\hat{p}_y, z\hat{p}_x] + [y\hat{p}_x, z\hat{p}_x]$$

Again, by components

$$[x\hat{p}_y, x\hat{p}_z] = [x\hat{p}_y, x]\hat{p}_z + x[x\hat{p}_y, \hat{p}_z] = -[x, x]\hat{p}_y\hat{p}_z - x[x, \hat{p}_y]\hat{p}_z - x[\hat{p}_z, x]\hat{p}_y - x^2[\hat{p}_z, \hat{p}_y] = 0$$

$$-[y\hat{p}_x, x\hat{p}_z] = -[y\hat{p}_x, x]\hat{p}_z - x[y\hat{p}_x, \hat{p}_z] = [x, y]\hat{p}_x\hat{p}_z + y[x, \hat{p}_x]\hat{p}_z + x[\hat{p}_z, y]\hat{p}_x + xy[\hat{p}_z, \hat{p}_x] = i\hbar y\hat{p}_z$$

$$-[x\hat{p}_y, z\hat{p}_x] = -[x\hat{p}_y, z]\hat{p}_x - z[x\hat{p}_y, \hat{p}_x] = [z, x]\hat{p}_y\hat{p}_x + x[z, \hat{p}_y]\hat{p}_x + z[\hat{p}_x, x]\hat{p}_y + zx[\hat{p}_x, \hat{p}_z] = -i\hbar z\hat{p}_y$$

$$[y\hat{p}_x, z\hat{p}_x] = [y\hat{p}_x, z]\hat{p}_x + z[y\hat{p}_x, \hat{p}_x] = -[z, y]\hat{p}_x^2 - y[z, \hat{p}_x]\hat{p}_x - z[\hat{p}_x, y]\hat{p}_x - zy[\hat{p}_x, \hat{p}_x] = 0$$

Hence

$$[\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x$$

Finally

$$[\hat{L}_z, \hat{L}_x] = [x\hat{p}_y - y\hat{p}_x, y\hat{p}_z - z\hat{p}_y] = [x\hat{p}_y, y\hat{p}_z - z\hat{p}_y] - [y\hat{p}_x, y\hat{p}_z - z\hat{p}_y] =$$
$$= [y\hat{p}_z - z\hat{p}_y, y\hat{p}_x] - [y\hat{p}_z - z\hat{p}_y, x\hat{p}_y] = [y\hat{p}_z, y\hat{p}_x] - [z\hat{p}_y, y\hat{p}_x] - [y\hat{p}_z, x\hat{p}_y] + [z\hat{p}_y, x\hat{p}_y]$$

So

$$[y\hat{p}_z, y\hat{p}_x] = [y\hat{p}_z, y]\hat{p}_x + y[y\hat{p}_z, \hat{p}_x] = -[y, y]\hat{p}_z\hat{p}_x - y[y, \hat{p}_z]\hat{p}_x - y[\hat{p}_x, y]\hat{p}_z - y^2[\hat{p}_x, \hat{p}_z] = 0$$

$$-[z\hat{p}_y, y\hat{p}_x] = -[z\hat{p}_y, y]\hat{p}_x - y[z\hat{p}_y, \hat{p}_x] = [y, z]\hat{p}_y\hat{p}_x + z[y, \hat{p}_y]\hat{p}_x + y[\hat{p}_x, z]\hat{p}_y + yz[\hat{p}_x, \hat{p}_y] = i\hbar z\hat{p}_x$$

$$-[y\hat{p}_z, x\hat{p}_y] = -[y\hat{p}_z, x]\hat{p}_y - x[y\hat{p}_z, \hat{p}_y] = [x, y]\hat{p}_z\hat{p}_y + y[x, \hat{p}_z]\hat{p}_y + x[\hat{p}_y, y]\hat{p}_z + xy[\hat{p}_y, \hat{p}_z] = -i\hbar x\hat{p}_z$$

$$[z\hat{p}_y, x\hat{p}_y] = [z\hat{p}_y, x]\hat{p}_y + x[z\hat{p}_y, \hat{p}_y] = -[x, z]\hat{p}_y^2 - z[x, \hat{p}_y]\hat{p}_y - x[\hat{p}_y, z]\hat{p}_y - xz[\hat{p}_y, \hat{p}_y] = 0$$

So

$$[\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y$$

Summarizing, the commutators are cyclic as follows

$$[\hat{L}_x, \hat{L}_y] = i\hbar L_z \tag{25}$$

$$[\hat{L}_y, \hat{L}_z] = i\hbar L_x \tag{26}$$

$$[\hat{L}_z, \hat{L}_x] = i\hbar L_y \tag{27}$$

Therefore, we are not able to know any two components of the angular momentum at the same time. But what about the magnitude of the angular momentum. Again, we need to find commutators of the magnitude operator $\hat{L}^2$ with operator for each component of angular momentum. Starting with x

$$[\hat{L}^2, \hat{L}_x] = [\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2, \hat{L}_x] = [\hat{L}_x^2, \hat{L}_x] + [\hat{L}_y^2, \hat{L}_x] + [\hat{L}_z^2, \hat{L}_x]$$

Commutator of any operator with the power of the same operator is always zero, as the operators in the operator sequence are identical and swapping them does not change the form of the operator sequence. By operator sequence I mean the product of the two operators in the commutator.
Therefore $[\hat{L}_x^2, \hat{L}_x] = 0$.
Then

$$[\hat{L}^2, \hat{L}_x] = [\hat{L}_y^2, \hat{L}_x] + [\hat{L}_z^2, \hat{L}_x] = -[\hat{L}_x, \hat{L}_y]\hat{L}_y - \hat{L}_y[\hat{L}_x, \hat{L}_y] - [\hat{L}_x, \hat{L}_z]\hat{L}_z - \hat{L}_z[\hat{L}_x, \hat{L}_z] =$$

$$= -i\hbar\hat{L}_z\hat{L}_y - i\hbar\hat{L}_y\hat{L}_z + i\hbar\hat{L}_y\hat{L}_z + i\hbar\hat{L}_z\hat{L}_y = 0$$

Similarly

$$[\hat{L}^2, \hat{L}_y] = -[\hat{L}_y, \hat{L}_x^2] - [\hat{L}_y, \hat{L}_y^2] - [\hat{L}_y, \hat{L}_z^2] = -[\hat{L}_y, \hat{L}_x]\hat{L}_x - \hat{L}_x[\hat{L}_y, \hat{L}_x] - [\hat{L}_y, \hat{L}_z]\hat{L}_z - \hat{L}_z[\hat{L}_y, \hat{L}_z] =$$

$$= i\hbar\hat{L}_z\hat{L}_x + i\hbar\hat{L}_x\hat{L}_z - i\hbar\hat{L}_x\hat{L}_z - i\hbar\hat{L}_z\hat{L}_x = 0$$

$$[\hat{L}^2, \hat{L}_z] = -[\hat{L}_z, \hat{L}_x^2] - [\hat{L}_z, \hat{L}_y^2] = -[\hat{L}_z, \hat{L}_x]\hat{L}_x - \hat{L}_x[\hat{L}_z, \hat{L}_x] - [\hat{L}_z, \hat{L}_y]\hat{L}_y - \hat{L}_y[\hat{L}_z, \hat{L}_y] =$$

$$= -i\hbar\hat{L}_y\hat{L}_x - i\hbar\hat{L}_x\hat{L}_y + i\hbar\hat{L}_x\hat{L}_y + i\hbar\hat{L}_y\hat{L}_x = 0$$

This means that we can know both one component of the angular momentum and its magnitude at the same time. It also means that $\hat{L}^2$ and $\hat{L}_x$ (or any other single component) are compatible operators, meaning that they share a common set of eigenfunctions. These eigenfunctions are called the spherical harmonics, and they are discovered later in their own subsection.

2.4.2   Angular Momentum in Spherical Polar Coordinates

We can change the coordinate system to find cartesian components of the angular momentum in terms of derivatives and multiples of spherical coordinates. This is especially useful if we have a wavefunction given in spherical coordinates.
The spherical polar coordinate system is given by a following series of transforms

$$x = r\sin\theta\cos\phi$$

$$y = r\sin\theta\sin\phi$$

$$z = r\cos\theta$$

and inverse transforms

$$r = \sqrt{x^2 + y^2 + z^2}$$

$$\theta = \cos^{-1}\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right)$$

$$\phi = \tan^{-1}\left(\frac{y}{x}\right)$$

Now, we need to transform the momentum operators. Starting with $\hat{L}_x$

$$\hat{L}_x = y\hat{p}_z - z\hat{p}_y = (-i\hbar)\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)$$

Here

$$\frac{\partial}{\partial z} = \frac{\partial r}{\partial z}\frac{\partial}{\partial r} + \frac{\partial\theta}{\partial z}\frac{\partial}{\partial\theta} + \frac{\partial\phi}{\partial z}\frac{\partial}{\partial\phi}$$

$$\frac{\partial r}{\partial z} = \frac{z}{r} = \cos\theta$$

$$\frac{\partial\theta}{\partial z} = -\frac{1}{\sqrt{1-\frac{z^2}{r^2}}}\left(\frac{r - z\frac{z}{r}}{r^2}\right) = \frac{z^2 - r^2}{r^3\sqrt{1-\frac{z^2}{r^2}}} = \frac{-(1-\cos^2\theta)}{r\sqrt{1-\cos^2\theta}} = \frac{-\sin\theta}{r}$$

$$\frac{\partial \phi}{\partial z} = 0$$

Also

$$\frac{\partial}{\partial y} = \frac{\partial r}{\partial y}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial y}\frac{\partial}{\partial \theta} + \frac{\partial \phi}{\partial y}\frac{\partial}{\partial \phi}$$

$$\frac{\partial r}{\partial y} = \frac{y}{r} = \sin\theta\sin\phi$$

$$\frac{\partial \theta}{\partial y} = -\frac{1}{\sqrt{1 - \frac{z^2}{r^2}}}\frac{-z}{r^2}\frac{y}{r} = \frac{\cos\theta\sin\phi}{r}$$

$$\frac{\partial \phi}{\partial y} = \frac{1}{1 + \frac{y^2}{x^2}}\frac{1}{x} = \frac{x}{y^2 + x^2} = \frac{r\sin\theta\cos\phi}{r^2\sin^2\theta(\sin^2\phi + \cos^2\phi)} = \frac{\cos\phi}{r\sin\theta}$$

So, putting the first term together

$$y\frac{\partial}{\partial z} = r\sin\theta\sin\phi\left(\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\right) = r\sin\theta\cos\theta\sin\phi\frac{\partial}{\partial r} - \sin^2\theta\sin\phi\frac{\partial}{\partial \theta}$$

And the second term together

$$-z\frac{\partial}{\partial y} = -r\cos\theta\left(\sin\theta\sin\phi\frac{\partial}{\partial r} + \frac{\cos\theta\sin\phi}{r}\frac{\partial}{\partial \theta} + \frac{\cos\phi}{r\sin\theta}\frac{\partial}{\partial \phi}\right) =$$

$$= -r\sin\theta\cos\theta\sin\phi\frac{\partial}{\partial r} - \cos^2\theta\sin\phi\frac{\partial}{\partial \theta} - \frac{\cos\theta\cos\phi}{\sin\theta}\frac{\partial}{\partial \phi}$$

Putting these two together

$$y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y} = -\sin\phi(\sin^2\theta + \cos^2\theta)\frac{\partial}{\partial \theta} - \frac{\cos\theta\cos\phi}{\sin\theta}\frac{\partial}{\partial \phi} = -\sin\phi\frac{\partial}{\partial \theta} - \frac{\cos\theta\cos\phi}{\sin\theta}\frac{\partial}{\partial \phi}$$

Hence

$$\hat{L}_x = (-i\hbar)\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) = i\hbar\sin\phi\frac{\partial}{\partial \theta} + i\hbar\cot\theta\cos\phi\frac{\partial}{\partial \phi}$$

Similarly for $\hat{L}_y$

$$\hat{L}_y = (-i\hbar)\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)$$

$$\frac{\partial}{\partial x} = \frac{\partial r}{\partial x}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial x}\frac{\partial}{\partial \theta} + \frac{\partial \phi}{\partial x}\frac{\partial}{\partial \phi}$$

$$\frac{\partial r}{\partial x} = \frac{x}{r} = \sin\theta\cos\phi$$

$$\frac{\partial \theta}{\partial x} = -\frac{1}{\sqrt{1 - \cos^2\theta}}\frac{z}{r^2}\frac{-x}{r} = \frac{\cos\theta\cos\phi}{r}$$

$$\frac{\partial \phi}{\partial x} = \frac{1}{1 + \frac{y^2}{x^2}}\frac{-y}{x^2} = -\frac{r\sin\theta\sin\phi}{r^2\sin^2\theta(\cos^2\phi + \sin^2\phi)} = -\frac{\sin\phi}{r\sin\theta}$$

$$z\frac{\partial}{\partial x} = r\cos\theta\left(\sin\theta\cos\phi\frac{\partial}{\partial r} + \frac{\cos\theta\cos\phi}{r}\frac{\partial}{\partial \theta} - \frac{\sin\phi}{r\sin\theta}\frac{\partial}{\partial \phi}\right) =$$

$$= r\sin\theta\cos\theta\cos\phi\frac{\partial}{\partial r} + \cos^2\theta\cos\phi\frac{\partial}{\partial \theta} - \frac{\cos\theta\sin\phi}{\sin\theta}\frac{\partial}{\partial \phi}$$

$$-x\frac{\partial}{\partial z} = -r\sin\theta\cos\phi\left(\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\right) = -r\sin\theta\cos\theta\cos\phi\frac{\partial}{\partial r} + \sin^2\theta\cos\phi\frac{\partial}{\partial \theta}$$

$$z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z} = \cos\phi\frac{\partial}{\partial \theta} - \frac{\cos\theta\sin\phi}{\sin\theta}\frac{\partial}{\partial \phi}$$

$$\hat{L}_y = (-i\hbar)\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) = -i\hbar\cos\phi\frac{\partial}{\partial \theta} + i\hbar\cot\theta\sin\phi\frac{\partial}{\partial \phi}$$

Finally, for $\hat{L}_z$

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

$$x\frac{\partial}{\partial y} = r\sin\theta\cos\phi\left(\sin\theta\sin\phi\frac{\partial}{\partial r} + \frac{\cos\theta\sin\phi}{r}\frac{\partial}{\partial\theta} + \frac{\cos\phi}{r\sin\theta}\frac{\partial}{\partial\phi}\right) =$$

$$= r\sin^2\theta\sin\phi\cos\phi\frac{\partial}{\partial r} + \sin\theta\cos\theta\sin\phi\cos\phi\frac{\partial}{\partial\theta} + \cos^2\phi\frac{\partial}{\partial\phi}$$

$$-y\frac{\partial}{\partial x} = -r\sin\theta\sin\phi\left(\sin\theta\cos\phi\frac{\partial}{\partial r} + \frac{\cos\theta\cos\phi}{r}\frac{\partial}{\partial\theta} - \frac{\sin\phi}{r\sin\theta}\frac{\partial}{\partial\phi}\right) =$$

$$= -r\sin^2\theta\sin\phi\cos\phi\frac{\partial}{\partial r} - \sin\theta\cos\theta\sin\phi\cos\phi\frac{\partial}{\partial\theta} + \sin^2\phi\frac{\partial}{\partial\phi}$$

Hence

$$x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} = (\cos^2\phi + \sin^2\phi)\frac{\partial}{\partial\phi} = \frac{\partial}{\partial\phi}$$

And

$$\hat{L}_z = (-i\hbar)\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) = -i\hbar\frac{\partial}{\partial\phi}$$

In summary

$$\hat{L}_x = i\hbar\sin\phi\frac{\partial}{\partial\theta} + i\hbar\cot\theta\cos\phi\frac{\partial}{\partial\phi} \tag{28}$$

$$\hat{L}_y = -i\hbar\cos\phi\frac{\partial}{\partial\theta} + i\hbar\cot\theta\sin\phi\frac{\partial}{\partial\phi} \tag{29}$$

$$\hat{L}_z = -i\hbar\frac{\partial}{\partial\phi} \tag{30}$$

We can see that it is much easier to obtain the z component of the angular momentum from explicitly written wavefunction $\Psi(r, \theta, \phi)$. Hence, if we will consider some one component of angular momentum, we will usually mean the z component, as it is much easier to obtain in spherical polar coordinates.
Another important property to notice is that none of the operators depends in any way on r. This means that the eigenfunctions of these operators (spherical harmonics) are only dependent on θ and ϕ, and multiplying them by some function of radius is the same as multiplication by a scalar constant.
From these forms, we can calculate the form of $\hat{L}^2$ in spherical polar coordinates.

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$$

$$\hat{L}_x^2 = -\hbar^2\left(\sin^2\phi\frac{\partial^2}{\partial\theta^2} + \sin\phi\frac{\partial}{\partial\theta}\cot\theta\cos\phi\frac{\partial}{\partial\phi} + \cot\theta\cos\phi\frac{\partial}{\partial\phi}\sin\phi\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi}\cot\theta\cos\phi\frac{\partial}{\partial\phi}\right) =$$

$$= -\hbar^2\left(\sin^2\phi\frac{\partial^2}{\partial\theta^2} - \sin\phi\cos\phi\frac{1}{\sin^2\theta}\frac{\partial}{\partial\phi} + \sin\phi\cos\phi\cot\theta\frac{\partial^2}{\partial\theta\partial\phi} + \cot\theta\cos^2\phi\frac{\partial}{\partial\theta}+\right.$$

$$\left.+ \sin\phi\cos\phi\cot\theta\frac{\partial^2}{\partial\theta\partial\phi} - \cot^2\theta\sin\phi\cos\phi\frac{\partial}{\partial\phi} + \cot^2\theta\cos^2\phi\frac{\partial^2}{\partial\phi^2}\right) =$$

$$= -\hbar^2\left(\sin^2\phi\frac{\partial^2}{\partial\theta^2} - \sin\phi\cos\phi\left(\frac{1}{\sin^2\theta} + \cot^2\theta\right)\frac{\partial}{\partial\phi} + 2\sin\phi\cos\phi\cot\theta\frac{\partial^2}{\partial\theta\partial\phi}+\right.$$

$$\left.+ \cot\theta\cos^2\phi\frac{\partial}{\partial\theta} + \cot^2\theta\cos^2\phi\frac{\partial^2}{\partial\phi^2}\right)$$

Similarly

$$\hat{L}_y^2 = -\hbar^2\left(\cos^2\phi\frac{\partial^2}{\partial\theta^2} - \cos\phi\frac{\partial}{\partial\theta}\cot\theta\sin\phi\frac{\partial}{\partial\phi} - \cot\theta\sin\phi\frac{\partial}{\partial\phi}\cos\phi\frac{\partial}{\partial\theta} + \cot\theta\sin\phi\frac{\partial}{\partial\phi}\cot\theta\sin\phi\frac{\partial}{\partial\phi}\right) =$$

$$= -\hbar^2 \left(\cos^2 \phi \frac{\partial^2}{\partial \theta^2} + \sin \phi \cos \phi \frac{1}{\sin^2 \theta} \frac{\partial}{\partial \phi} - \sin \phi \cos \phi \cot \theta \frac{\partial^2}{\partial \theta \partial \phi} + \cot \theta \sin^2 \phi \frac{\partial}{\partial \theta} - \sin \phi \cos \phi \cot \theta \frac{\partial^2}{\partial \theta \partial \phi} + \right.$$

$$\left. + \cot^2 \theta \sin \phi \cos \phi \frac{\partial}{\partial \phi} + \cot^2 \theta \sin^2 \phi \frac{\partial^2}{\partial \phi^2} \right) =$$

$$= -\hbar^2 \left(\cos^2 \phi \frac{\partial^2}{\partial \theta^2} - 2 \cot \theta \sin \phi \cos \phi \frac{\partial^2}{\partial \theta \partial \phi} + \sin \phi \cos \phi \left(\frac{1}{\sin^2 \theta} + \cot^2 \theta \right) \frac{\partial}{\partial \phi} + \right.$$

$$\left. + \cot \theta \sin^2 \phi \frac{\partial}{\partial \theta} + \cot^2 \theta \sin^2 \phi \frac{\partial^2}{\partial^2 \phi} \right)$$

Adding together $\hat{L}_x^2$ and $\hat{L}_y^2$ leads to major simplification

$$\hat{L}_x^2 + \hat{L}_y^2 = -\hbar^2 \left((\sin^2 \phi + \cos^2 \phi) \frac{\partial^2}{\partial \theta^2} + \cot \theta (\cos^2 \phi + \sin^2 \phi) \frac{\partial}{\partial \theta} + \cot^2 \theta (\cos^2 \phi + \sin^2 \phi) \frac{\partial^2}{\partial \phi^2} \right) =$$

$$= -\hbar^2 \left(\frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \cot^2 \theta \frac{\partial^2}{\partial \phi^2} \right) = -\hbar^2 \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \cot^2 \theta \frac{\partial^2}{\partial \phi^2} \right)$$

Finally,

$$\hat{L}_z^2 = -\hbar^2 \frac{\partial^2}{\partial \phi^2}$$

And therefore

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 = -\hbar^2 \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + (\cot^2 \theta + 1) \frac{\partial^2}{\partial \phi^2} \right)$$

$$\hat{L}^2 = -\hbar^2 \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right)$$

2.4.3 Commutation with Hamiltonian Operator

Now, we return to cartesian coordinates and discover the commutators of hamiltonian operator and angular momentum operators. Starting from $\hat{L}_x$ and using full form of hamiltonian (3D)

$$\hat{H} = -\frac{\hbar^2}{2m} \nabla^2 + V(x, y, z) = \frac{1}{2m} \left(\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2 \right) + V(x, y, z)$$

$$[\hat{L}_x, \hat{H}] = [y\hat{p}_z - z\hat{p}_y, \frac{1}{2m} \left(\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2 \right) + V] =$$

$$= [y\hat{p}_z, \frac{\hat{p}_x^2}{2m} + \frac{\hat{p}_y^2}{2m} + \frac{\hat{p}_z^2}{2m} + V] - [z\hat{p}_y, \frac{\hat{p}_x^2}{2m} + \frac{\hat{p}_y^2}{2m} + \frac{\hat{p}_z^2}{2m} + V] =$$

$$= [\frac{\hat{p}_x^2}{2m} + \frac{\hat{p}_y^2}{2m} + \frac{\hat{p}_z^2}{2m} + V, z\hat{p}_y] - [\frac{\hat{p}_x^2}{2m} + \frac{\hat{p}_y^2}{2m} + \frac{\hat{p}_z^2}{2m} + V, y\hat{p}_z] =$$

$$= \frac{1}{2m} \left([\hat{p}_x^2, z\hat{p}_y] + [\hat{p}_y^2, z\hat{p}_y] + [\hat{p}_z^2, z\hat{p}_y] - [\hat{p}_x^2, y\hat{p}_z] - [\hat{p}_y^2, y\hat{p}_z] - [\hat{p}_z^2, y\hat{p}_z] \right) + [V, z\hat{p}_y] - [V, y\hat{p}_z] =$$

$$= \frac{1}{2m} \left([\hat{p}_z^2, z\hat{p}_y] - [\hat{p}_y^2, y\hat{p}_z] \right) + [V, z]\hat{p}_y + z[V, \hat{p}_y] - [V, y]\hat{p}_z - y[V, \hat{p}_z] =$$

$$= \frac{1}{2m} \left([\hat{p}_z^2, z]\hat{p}_y + z[\hat{p}_z^2, \hat{p}_y] - [\hat{p}_y^2, y]\hat{p}_z - y[\hat{p}_y^2, \hat{p}_z] \right) + z[V, \hat{p}_y] - y[V, \hat{p}_z] =$$

$$= \frac{1}{2m} \left([y, \hat{p}_y^2]\hat{p}_z - [z, \hat{p}_z^2]\hat{p}_y \right) + z[V, \hat{p}_y] - y[V, \hat{p}_z] =$$

$$= \frac{1}{2m} \left([y, \hat{p}_y]\hat{p}_y\hat{p}_z + \hat{p}_y[y, \hat{p}_y]\hat{p}_z - [z, \hat{p}_z]\hat{p}_z\hat{p}_y - \hat{p}_z[z, \hat{p}_z]\hat{p}_y \right) + z[V, \hat{p}_y] - y[V, \hat{p}_z]$$

Since momentum operators commute and $[x_i, \hat{p}_{x_i}] = i\hbar$ is a constant for every $\hat{x}_i$, we have

$$[\hat{L}_x, \hat{H}] = \frac{i\hbar}{2m} \left(\hat{p}_y\hat{p}_z + \hat{p}_y\hat{p}_z - \hat{p}_z\hat{p}_y - \hat{p}_z\hat{p}_y \right) + y[\hat{p}_z, V] - z[\hat{p}_y, V] = y[\hat{p}_z, V] - z[\hat{p}_y, V]$$

Similarly

$$[\hat{L}_y, \hat{H}] = [\frac{1}{2m}(\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2) + V, x\hat{p}_z] - [\frac{1}{2m}(\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2) + V, z\hat{p}_x] =$$

$$= \frac{1}{2m}\left([\hat{p}_x^2, x\hat{p}_z] + [\hat{p}_y^2, x\hat{p}_z] + [\hat{p}_z^2, x\hat{p}_z] - [\hat{p}_x^2, z\hat{p}_x] - [\hat{p}_y^2, z\hat{p}_x] - [\hat{p}_z^2, z\hat{p}_x]\right) + z[\hat{p}_x, V] - x[\hat{p}_z, V] =$$

$$= \frac{1}{2m}\left([\hat{p}_x^2, x]\hat{p}_z + x[\hat{p}_z^2, \hat{p}_z] - [\hat{p}_z^2, z]\hat{p}_x - z[\hat{p}_z^2, \hat{p}_x]\right) + z[\hat{p}_x, V] - x[\hat{p}_z, V] =$$

$$= \frac{1}{2m}\left([z, \hat{p}_z]\hat{p}_z\hat{p}_x + \hat{p}_z[z, \hat{p}_z]\hat{p}_x - [x, \hat{p}_x]\hat{p}_x\hat{p}_z - \hat{p}_x[x, \hat{p}_x]\hat{p}_z\right) + z[\hat{p}_x, V] - x[\hat{p}_z, V] =$$

$$= \frac{i\hbar}{2m}\left(\hat{p}_z\hat{p}_x + \hat{p}_z\hat{p}_x - \hat{p}_x\hat{p}_z - \hat{p}_x\hat{p}_z\right) + z[\hat{p}_x, V] - x[\hat{p}_z, V] = z[\hat{p}_x, V] - x[\hat{p}_z, V]$$

And

$$[\hat{L}_z, \hat{H}] = [\frac{1}{2m}(\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2) + V, y\hat{p}_x] - [\frac{1}{2m}(\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2) + V, x\hat{p}_y] =$$

$$= \frac{1}{2m}\left([\hat{p}_x^2, y\hat{p}_x] + [\hat{p}_y^2, y\hat{p}_x] + [\hat{p}_z^2, y\hat{p}_x] - [\hat{p}_x^2, x\hat{p}_y] - [\hat{p}_y^2, x\hat{p}_y] - [\hat{p}_z^2, x\hat{p}_y]\right) + x[\hat{p}_y, V] - y[\hat{p}_x, V] =$$

$$= \frac{1}{2m}\left([\hat{p}_y^2, y]\hat{p}_x + y[\hat{p}_y^2, \hat{p}_x] - [\hat{p}_x^2, x]\hat{p}_y - x[\hat{p}_x^2, \hat{p}_y]\right) + x[\hat{p}_y, V] - y[\hat{p}_x, V] =$$

$$= \frac{1}{2m}\left([x, \hat{p}_x]\hat{p}_x\hat{p}_y + \hat{p}_x[x, \hat{p}_x]\hat{p}_y - [y, \hat{p}_y]\hat{p}_y\hat{p}_x - \hat{p}_y[y, \hat{p}_y]\hat{p}_x\right) + x[\hat{p}_y, V] - y[\hat{p}_x, V] =$$

$$= \frac{i\hbar}{2m}\left(\hat{p}_x\hat{p}_y + \hat{p}_x\hat{p}_y - \hat{p}_y\hat{p}_x - \hat{p}_y\hat{p}_x\right) + x[\hat{p}_y, V] - y[\hat{p}_x, V] = x[\hat{p}_y, V] - y[\hat{p}_x, V]$$

In summary

$$[\hat{L}_x, \hat{H}] = y[\hat{p}_z, V] - z[\hat{p}_y, V]$$
$$[\hat{L}_y, \hat{H}] = z[\hat{p}_x, V] - x[\hat{p}_z, V]$$
$$[\hat{L}_z, \hat{H}] = x[\hat{p}_y, V] - y[\hat{p}_x, V]$$

To calculate the commutators involving the potential, we use

$$[\hat{p}_{x_i}, V] = (-i\hbar)\left(\frac{\partial}{\partial x_i}V - V\frac{\partial}{\partial x_i}\right) = (-i\hbar)\frac{\partial V}{\partial x_i}$$

Hence

$$[\hat{L}_x, \hat{H}] = y(-i\hbar)\frac{\partial V}{\partial z} - z(-i\hbar)\frac{\partial V}{\partial y} = -i\hbar(\vec{r} \times (\nabla V))_x \tag{31}$$

$$[\hat{L}_y, \hat{H}] = -i\hbar(\vec{r} \times (\nabla V))_y \tag{32}$$

$$[\hat{L}_z, \hat{H}] = -i\hbar(\vec{r} \times (\nabla V))_z \tag{33}$$

where $\vec{r}$ is the position vector and ∇V is the gradient of potential.
Interesting is a special case when potential is radial, i.e. $V = V(r)$, where r is radius in spherical polar coordinates. Then

$$\nabla V = \frac{\partial V}{\partial r}\hat{e}_r$$

where $\hat{e}_r$ is the unit vector in the direction of position vector $\vec{r}$.
But in this case

$$\vec{r} \times (\nabla V) = \frac{\partial V}{\partial r}r(\hat{e}_r \times \hat{e}_r) = \vec{0}$$

where $\vec{0}$ is the zero vector. This means that

$$[\hat{L}_x, \hat{H}] = [\hat{L}_y, \hat{H}] = [\hat{L}_z, \hat{H}] = 0$$

and thus also

$$[\hat{L}^2, \hat{H}] = [\hat{L}_x^2, \hat{H}] + [\hat{L}_y^2, \hat{H}] + [\hat{L}_z^2, \hat{H}] =$$

$$= [\hat{L}_x, \hat{H}]\hat{L}_x + \hat{L}_x[\hat{L}_x, \hat{H}] + [\hat{L}_y, \hat{H}]\hat{L}_y + \hat{L}_y[\hat{L}_y, \hat{H}] + [\hat{L}_z, \hat{H}]\hat{L}_z + \hat{L}_z[\hat{L}_z, \hat{H}] = 0$$

Therefore, for central potential, hamiltonian operator, angular momentum magnitude operator and one angular momentum component operator are all compatible operators and they have a common set of eigenfunctions. These eigenfunctions consist of spherical harmonics (which only depend on θ and ϕ), which are multiplied by some radial function $R(r)$. Therefore, the eigenstates of energy for radial potential can be written as

$$\Psi(r, \theta, \phi) = R(r)Y(\theta, \phi) \tag{34}$$

where $Y(\theta, \phi)$ is the spherical harmonic.

2.5   Spherical Harmonics

Spherical harmonics are the eigenfunctions of $\hat{L}^2$ and $\hat{L}_z$ operators (choosing component $\hat{L}_z$ because of its simple form in spherical coordinates). The eigenvalue equations are

$$\hat{L}_z Y = l_z Y$$

$$\hat{L}^2 Y = L^2 Y$$

where L is the magnitude of the angular momentum and l_z is the component of the angular momentum along the z direction.

2.5.1   Raising and Lowering Operators

Now consider two artificial operators

$$\hat{L}_+ = \hat{L}_x + i\hat{L}_y$$

$$\hat{L}_- = \hat{L}_x - i\hat{L}_y$$

To understand the properties of these operators, consider applying $\hat{L}_+$ to the $\hat{L}_z$ eigenvalue equation. Lets also denote the specific function which gives eigenvalue l_z as $Y = Y_{lm_l}$

$$\hat{L}_z Y_{lm_l} = l_z Y_{lm_l}$$

$$\hat{L}_+ \hat{L}_z Y_{lm_l} = l_z L_+ Y_{lm_l}$$

$$([\hat{L}_+, \hat{L}_z] + \hat{L}_z \hat{L}_+) Y_{lm_l} = l_z \hat{L}_+ Y_{lm_l}$$

Now we find the commutator

$$[\hat{L}_+, \hat{L}_z] = [\hat{L}_x + i\hat{L}_y, \hat{L}_z] = [\hat{L}_x, \hat{L}_z] + i[\hat{L}_y, \hat{L}_z] = -i\hbar\hat{L}_y + i(i\hbar)\hat{L}_x = (-\hbar)(\hat{L}_x + i\hat{L}_y) = -\hbar\hat{L}_+$$

Hence the equation becomes

$$(-\hbar\hat{L}_+ + \hat{L}_z \hat{L}_+) Y_{lm_l} = l_z \hat{L}_+ Y_{lm_l}$$

$$(\hat{L}_z - \hbar)(\hat{L}_+ Y_{lm_l}) = l_z(\hat{L}_+ Y_{lm_l})$$

$$\hat{L}_z(\hat{L}_+ Y_{lm_l}) = (l_z + \hbar)(\hat{L}_+ Y_{lm_l})$$

This means that $\hat{L}_+ Y_{lm_l}$ is also an eigenvalue equation, with eigenvalue $l_z + \hbar$. We can label

$$\hat{L}_+ Y_{lm_l} = Y_{l(m_l+1)}$$

Similarly

$$\hat{L}_- \hat{L}_z Y_{lm_l} = l_z \hat{L}_- Y_{lm_l}$$

$$[\hat{L}_-, \hat{L}_z] = [\hat{L}_x, \hat{L}_z] - i[\hat{L}_y, \hat{L}_z] = -i\hbar\hat{L}_y - i(i\hbar)\hat{L}_x = \hbar(\hat{L}_x - i\hat{L}_y) = \hbar\hat{L}_-$$

$$\hat{L}_- \hat{L}_z Y_{lm_l} = ([\hat{L}_-, \hat{L}_z] + \hat{L}_z \hat{L}_-) Y_{lm_l} = l_z \hat{L}_- Y_{lm_l}$$

$$(\hat{L}_z + \hbar)(\hat{L}_- Y_{lm_l}) = l_z \hat{L}_- Y_{lm_l}$$

Labeling $\hat{L}_- Y_{lm_l} = Y_{l(m_l-1)}$

$$(\hat{L}_z + \hbar) Y_{l(m_l-1)} = l_z Y_{l(m_l-1)}$$

$$\hat{L}_z Y_{l(m_l-1)} = (l_z - \hbar) Y_{l(m_l-1)}$$

So, $Y_{l(m_l-1)}$ is also an eigenfunction of $\hat{L}_z$ and has eigenvalue $l_z - \hbar$.
Furthemore, lets apply this operator to the other eigenvalue equation

$$\hat{L}_+ \hat{L}^2 Y_{lm_l} = L^2 \hat{L}_+ Y_{lm_l}$$

$$([\hat{L}_+, \hat{L}^2] + \hat{L}^2 \hat{L}_+) Y_{lm_l} = L^2 Y_{l(m_l+1)}$$

$$[\hat{L}_+, \hat{L}^2] = [\hat{L}_x, \hat{L}^2] + i[\hat{L}_y, \hat{L}^2] = 0$$

So

$$\hat{L}^2 Y_{l(m_l+1)} = L^2 Y_{l(m_l+1)}$$

which means that $Y_{l(m_l+1)}$ is also an eigenvalue of the $\hat{L}^2$ operator with the same eigenvalue as Y_{lm_l} (hence the second index l in Y, which will be used to distinguish between different total angular momenta L). Similarly

$$[\hat{L}_-, \hat{L}^2] = [\hat{L}_x, \hat{L}^2] - i[\hat{L}_y, \hat{L}^2] = 0$$

$$\hat{L}_- \hat{L}^2 Y_{lm_l} = L^2 Y_{l(m_l-1)}$$

$$\hat{L}^2 Y_{l(m_l-1)} = L^2 Y_{l(m_l-1)}$$

It should be noted though that operating with $\hat{L}_-$ does not produce the function properly normalized - the function is multiplied by some constant of dimensions of angular momentum. Similarly for $\hat{L}_+$, the eigenfunction $Y_{l(m_l+1)}$ is multiplied by some constant of dimensions of angular momentum. Because of this, we cannot write that operating consecutively with $\hat{L}_+$ and $\hat{L}_-$ produces the original function, i. e.

$$\hat{L}_+ \hat{L}_- Y_{lm_l} \neq Y_{lm_l}$$

For the eigenvalue of $\hat{L}_z$ to make sense, its absolute value must be less than or equal to L - maximum value of component along z direction is the value of magnitude of the angular momentum. This means

$$|l_z| \leq L$$

$$-L \leq l_z \leq L$$

This means that there exists some maximum and minimum $Y_{lm_{l,max}}$ and $Y_{lm_{l,min}}$ for which we can no longer increase/decrease the value of l_z for any other consecutive application of $\hat{L}_\pm$. To ensure this, we set

$$\hat{L}_+ Y_{lm_{l,max}} = 0$$

$$\hat{L}_- Y_{lm_{l,min}} = 0$$

Applying $\hat{L}_-$ to both sides of the first equation gives us

$$\hat{L}_- \hat{L}_+ Y_{lm_{l,max}} = 0$$

$$([\hat{L}_-, \hat{L}_+] + \hat{L}_+ \hat{L}_-) Y_{lm_{l,max}} = 0$$

We now need to find commutator

$$[\hat{L}_-, \hat{L}_+] = [\hat{L}_x - i\hat{L}_y, \hat{L}_x + i\hat{L}_y] = [\hat{L}_x, \hat{L}_x + i\hat{L}_y] - i[\hat{L}_y, \hat{L}_x + \hat{L}_y] = i[\hat{L}_x + i\hat{L}_y, \hat{L}_y] - [\hat{L}_x + i\hat{L}_y, \hat{L}_x] =$$

$$= i[\hat{L}_x, \hat{L}_y] - [\hat{L}_y, \hat{L}_y] - [\hat{L}_x, \hat{L}_x] - i[\hat{L}_y, \hat{L}_x] = -\hbar\hat{L}_z + i[\hat{L}_x, \hat{L}_y] = -\hbar\hat{L}_z - \hbar\hat{L}_z = -2\hbar\hat{L}_z$$

Hence

$$(-2\hbar\hat{L}_z + \hat{L}_+ \hat{L}_-) Y_{lm_{l,max}} = 0$$

$$\hat{L}_+ \hat{L}_- Y_{lm_{l,max}} = 2\hbar\hat{L}_z Y_{lm_{l,max}}$$

Applying the operators explicitly

$$\hat{L}_+ \hat{L}_- = (\hat{L}_x + i\hat{L}_y)(\hat{L}_x - i\hat{L}_y) = \hat{L}_x^2 - i\hat{L}_x\hat{L}_y + i\hat{L}_y\hat{L}_x + \hat{L}_y^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 - \hat{L}_z^2 - i(\hat{L}_x, \hat{L}_y - \hat{L}_y) =$$

$$= \hat{L}^2 - \hat{L}_z^2 - i[\hat{L}_x, \hat{L}_y] = \hat{L}^2 - \hat{L}_z^2 - i(i\hbar)\hat{L}_z = \hat{L}^2 + \hbar\hat{L}_z - \hat{L}_z^2$$

And

$$(\hat{L}^2 + \hbar\hat{L}_z - \hat{L}_z^2) Y_{lm_{l,max}} = 2\hbar l_{z,max} Y_{lm_{l,max}}$$

Applying the eigenvalue equations

$$(L^2 + \hbar l_{z,max} - l_{z,max}^2) Y_{lm_{l,max}} = 2\hbar Y_{lm_{l,max}}$$

$$L^2 Y_{lm_{l,max}} = l_{z,max}(l_{z,max} + \hbar) Y_{lm_{l,max}}$$

For non trivial solutions, this means that

$$L^2 = l_{z,max}(l_{z,max} + \hbar) \tag{35}$$

Similarly, we can apply $\hat{L}_+$ to both sides for boundary condition for $Y_{lm_{l,min}}$

$$\hat{L}_+\hat{L}_-Y_{lm_{l,min}} = 0$$

$$([\hat{L}_+,\hat{L}_-] + \hat{L}_-\hat{L}_+)Y_{lm_{l,min}} = 0$$

$$(\hat{L}_-\hat{L}_+ - [\hat{L}_-,\hat{L}_+])Y_{lm_{l,min}} = 0$$

$$(\hat{L}_-\hat{L}_+ + 2\hbar\hat{L}_z)Y_{lm_{l,min}} = 0$$

$$\hat{L}_-\hat{L}_+Y_{lm_{l,min}} = -2\hbar\hat{L}_zY_{lm_{l,min}}$$

$$\hat{L}_-\hat{L}_+ = (\hat{L}_x - i\hat{L}_y)(\hat{L}_x + i\hat{L}_y) = \hat{L}_x^2 + \hat{L}_y^2 + i(\hat{L}_x\hat{L}_y - \hat{L}_y\hat{L}_x) = \hat{L}^2 - \hat{L}_z^2 + i(i\hbar)\hat{L}_z = \hat{L}^2 - \hat{L}_z^2 - \hbar\hat{L}_z$$

$$(\hat{L}^2 - \hat{L}_z^2 - \hbar\hat{L}_z)Y_{lm_{l,min}} = -2\hbar\hat{L}_zY_{lm_{l,min}}$$

$$(L^2 - l_{z,min}^2 - \hbar l_{z,min})Y_{lm_{l,min}} = -2\hbar l_{z,min}Y_{lm_{l,min}}$$

$$L^2Y_{lm_{l,min}} = l_{z,min}(l_{z,min} - \hbar)Y_{lm_{l,min}}$$

This means that

$$L^2 = l_{z,min}^2 - \hbar l_{z,min}$$

$$l_{z,max}^2 + \hbar l_{z,max} = l_{z,min}^2 - \hbar l_{z,min}$$

This equation has two possible solutions. One is trivial $l_{z,min} = l_{z,max} = 0$, other is non-trivial and states that

$$l_{z,max} = -l_{z,min}$$

Since the value of l_z always changes only by $\pm\hbar$ (using raising/lowering operators), the difference of the two extremal values must be an integer multiple of $\hbar$, i. e.

$$l_{z,max} - l_{z,min} = n\hbar$$

Together with the previous equation, we then derive that

$$l_{z,max} = \frac{n\hbar}{2} = -l_{z,min}$$

this means that the maximum value of component of angular momentum obtained by the operator is equal to a half integer if n is odd or full integer if n is even. This distinction has a close connection to the spin of particles.

In summary, we learned that the eigenvalues of total angular momentum operator $\hat{L}^2$ are L^2 such that

$$L = \hbar\sqrt{\frac{n}{2}\left(\frac{n}{2} + 1\right)} \tag{36}$$

where n is a positive integer. Usually, we denote $l = \frac{n}{2}$, where l is the l from the spherical harmonic lower index. Then $L = \hbar\sqrt{l(l+1)}$.

Also, the eigenvalues of the $\hat{L}_z$ operator are l_z such that

$$l_z = m_l\hbar \tag{37}$$

where m_l is an integer going from $-\frac{n}{2}$ to $\frac{n}{2}$.

2.5.2   Particle on a Sphere

If there is a particle confined to a sphere of constant radius but otherwise free to move, the potential for the particle is radial and in form of $-\delta(r)$.

Then, the energy eigenfunctions of the particle are spherical harmonics, since the hamiltonian in this case commutes with both $\hat{L}^2$ and $\hat{L}_z$. The general state is then a superposition of more spherical harmonics.

In order to be able to do a hamiltonian operation on such wavefunction, we need a explicit form of the wavefunction, and hence a explicit form of the spherical harmonic. For given total angular momentum $L = \hbar\sqrt{l(l+1)}$, the energy eigenvalue equation can be expressed as

$$\hat{H}Y_{lm_l} = E_{lm_l}Y_{lm_l}$$

The V is constant on the sphere, so it can be set to zero. Then, we are left with

$$-\frac{\hbar^2}{2m}\nabla^2 Y_{lm_l} = E_{lm_l} Y_{lm_l}$$

Where ∇^2 is the Laplacian operator. Since there is no dependence of Y_{lm_l} on radius r, the Laplacian operator is simply

$$\nabla^2 = \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2}{\partial\phi^2}$$

We see that this can be expressed in form of total angular momentum operator $\hat{L}^2$ as

$$\nabla^2 = \frac{-\hat{L}^2}{\hbar^2 r^2}$$

Hence

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 = \frac{\hat{L}^2}{2mr^2}$$

We see that for such a free particle on a sphere, only energy present is due to its angular momentum, which should not come as a surprise. Using our $\hat{L}^2$ eigenvalue equation

$$\hat{L}^2 Y_{lm_l} = \hbar^2 l(l+1) Y_{lm_l}$$

The energy eigenvalue equation becomes

$$\hat{H} Y_{lm_l} = E_{lm_l} Y_{lm_l}$$

$$\frac{\hat{L}^2}{2mr^2} Y_{lm_l} = E_{lm_l} Y_{lm_l}$$

$$\frac{\hbar^2 l(l+1)}{2mr^2} Y_{lm_l} = E_{lm_l} Y_{lm_l}$$

$$E_{lm_l} = \frac{\hbar^2 l(l+1)}{2mr^2}$$

And we thus found the energy of the particle.

2.5.3 Spin Angular Momentum

Spin is a property of a particle that behaves like an internal angular momentum of a particle. The mathematics and operators of spin angular momentum are taken from the operators of normal quantum angular momentum, i. e.

$$\hat{s}_z \Psi = m_l \hbar \Psi$$

where $m_l = \frac{n}{2}$ for fermions, and $m_l = n$ for bosons, where n is an integer (for photons one). Also

$$\hat{s}^2 \Psi = \hbar^2 s(s+1)\Psi$$

where $s = \frac{1}{2}$ for fermions (maximum m_l value) and $s = n$ for bosons. Slight complications arise from adding spins to the theory, because they cause state Ψ to have multiple possible values at the same points in space. This leads to Ψ not being a classical wavefunction but rather a spinor. More on this problematic is described in section on standard model and Dirac equation.

The spins are also cause for n in $m_{l,max} = \frac{n}{2}$ (derived maximum m_l for general central potential) to take any value, not only even values.

In fact, for fermions in central potential, the angular momentum along certain axis has to be a half-integer as well, because the fermion in central potential must still be a fermion.

3   Hydrogen Like Atoms

Hydrogen like atoms are atoms that consist of small, heavy, positively charged nucleaus and one electron orbiting said nucleus in a stationary state.

Since the nucleus is heavy, it does not move while the electron moves, which means that the potential acting on the electron is a central potential with origin in the nucleus. Because nucleus is so small, we assume it is point like.

As the potential is central, the energy eigenfunction is also an eigenfunction of the total angular momentum operator $\hat{L}^2$ and operator $\hat{L}_z$. This means that the energy eigenfunction is of type

$$\Psi_{nlm_l s}(r, \theta, \phi) = R_n(r) Y_{lm_l s}(\theta, \phi)$$

where we describe the function in terms of spherical polar coordinates and constant n is used to distinguish multiple possible wavefunctions and radial functions R. Spin number s is explicitly added because of historical reasons - l is not used in a form we derived thus far, but rather as a whole integer or zero that is smaller by $\frac{1}{2}$ than our predicted l. Our half-integer l is here usually denoted as j.

Since we would expect j to go as $j = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, ...$, l gere goes as $l = 0, 1, 2, ...$, and the half in total angular momentum is interpreted as adding the spin angular momentum of $\frac{1}{2}$ for the electron.

The energy eigenvalue equation is (coulombic potential in this case)

$$\hat{H}\Psi_{nlm_l s} = E_{nlm_l s}\Psi_n$$

Hence

$$\left(-\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{4\pi\epsilon_0 r}\right)\Psi_n = E_{nlm_l s}\Psi_n$$

Here, $\Psi_{nlm_l s}$ can depend on r. The laplacian is

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{\hat{L}^2}{\hbar^2 r^2}$$

Hence, since $\Psi_{nlm_l s}$ is an eigenfunction of $\hat{L}^2$ with eigenvalue being the total angular momentum

$$\hat{H}\Psi_{nlm_l s} = \left[-\frac{\hbar^2}{2m}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\hbar^2 j(j+1)}{2mr^2} - \frac{Ze^2}{4\pi\epsilon_0 r}\right]\Psi_{nlm_l s} = E_{nlm_l s}\Psi_{nlm_l s}$$

Since the operator on the left depends only on r, this equation has to apply only for radial equation as well, where partial derivatives become total derivatives

$$\left[-\frac{\hbar^2}{2mr^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right) + \frac{\hbar^2 j(j+1)}{2mr^2} - \frac{Ze^2}{4\pi\epsilon_0 r}\right]R_n(r) = E_{nlm_l s}R_n(r) \tag{38}$$

It turns out that the exact form of R_n also depends on j (or l, respectively). Hence, we should really name the function R_{nl}. The functions that solve this differential equation are of form

$$R_{nl} = C_{nl}P_l\left(\frac{r}{a_B}\right)e^{-\frac{Zr}{na_b}}$$

where P_l are the associated Laguerre polynomials, always up to the order of l, $a_b = 0.529 \times 10^{-10}$ m is the Bohr radius and C_{nl} is the normalisation constant.

The eigenvalues of the energy are surprisingly easy to obtain - they follow sequence

$$E_n = \frac{-Z^2 E_0}{n^2}$$

where $E_0 \approx 13.6$ eV is the ionisation energy of the electron when it is closest possible to the nucleus of hydrogen - so called Rydberg energy.

One important consequence of the eigenvalue equation which is not derived here is that number n is always bigger than zero, an integer and also always bigger than l.

3.1   Normalization

The wavefunction of the electron is normalized as

$$\int_V |\Psi_{nlm_l s}|^2 dV = 1$$

This can be expanded as

$$\int_0^\infty \int_0^\pi \int_0^{2\pi} |R(r)_{nl} Y_{lm_l}|^2 r^2 \sin\theta d\phi d\theta dr = 1$$

This can be then factorized as

$$\left(\int_0^\infty R_{nl}^* R_{nl} r^2 dr\right) \left(\int_0^\pi \int_0^{2\pi} Y_{lm_l}^* Y_{lm_l} \sin\theta d\phi d\theta\right) = 1$$

Usually, we normalize spherical harmonics separately. To normalize spherical harmonics, we require

$$\int_0^\pi \int_0^{2\pi} Y_{lm_l}^* Y_{lm_l} \sin\theta d\phi d\theta = 1$$

Which can be also written in terms of solid angle

$$\int_\Omega Y_{lm_l}^* Y_{lm_l} d\Omega = 1$$

where the integration goes over all space - full solid angle.
The normalization of the wavefunction than is simply

$$\int_0^\infty R_{nl}^* R_{nl} r^2 dr = 1$$

3.2   Zeeman Effect

So for, we could see that the Hamiltonian does not feature $\hat{L}_z$ operator - the energy of the electron is independent of the direction of the angular momentum (only depends on its magnitude). This however changes when the atom is inside a magnetic field. The energy of the electron iside the magnetic field is given by

$$E_m = -\vec{\mu}_e \cdot \vec{B}$$

where $\vec{B}$ is the magnetic induction vector and μ_e is the magnetic moment generated by the electron.
This moment is

$$\vec{\mu}_e = \vec{\mu} + \vec{\mu}_s$$

Where $\vec{\mu}$ is the magnetic moment due to electron orbiting around the nucleus and μ_s is the spin magnetic moment.
Since only the direction along $\vec{B}$ matters for energy determination, if we choose $\vec{B}$ to define z direction, $\hat{L}_z$ returns the angular momentum in z direction of the electron orbit.
The angular momentum of a point particle can be classically expressed as $\vec{L} = \vec{r} \times m\vec{v}$. The current due to this particle is then

$$\vec{I} = \frac{q}{T}\hat{v}$$

where T is the period of orbit of the particle, q is the charge of the particle and $\hat{v}$ is the unit vector in the direction of the velocity of the particle.
For a circular orbit, $\vec{r} \perp \vec{v}$, so

$$\vec{L} = mrv\hat{k}$$

if we choose the orbit in xy plane (and thus determine the z component of the angular momentum). Now, we are mostly interested in magnitudes of these vectors. Again, assuming circular orbit

$$I = \frac{q}{T} = \frac{q}{\frac{2\pi r}{v}} = \frac{qv}{2\pi r}$$

The magnetic moment will be then (A is the area swiped by the particle)

$$\vec{\mu} = IA\hat{k} = \frac{qv}{2\pi r}\pi r^2\hat{k} = \frac{qv}{2}r\hat{k} = \frac{q}{2}vr\hat{k} = \frac{q}{2m}\vec{L}$$

By correspondence principle, we then define the orbital magnetic moment operator as

$$\hat{\mu} = \frac{q}{2m}\hat{L}_z$$

which has, for our energy eigenstates, which are also eigenstates of $\hat{L}_z$, eigenvalues $\mu = \frac{q}{2m}\hbar m_l$. In case of electron

$$\mu = \frac{e}{2m_e}\hbar m_l$$

The above derivation was for a point particle in a magnetic field. But what about a spinning charged body in a magnetic field? For such body, we define the magnetic moment as sum of magnetic moments of the elements of the body, which are essentially the point particles, i.e. as an integral

$$\vec{\mu} = \int_V \frac{1}{2}\frac{dq}{dm}d\vec{L} = \frac{1}{2}\frac{q}{m}g\vec{L}$$

where g is the bodies gyromagnetic ratio. If the charged is spread evenly, g is a scalar, otherwise it is generally a tensor (similar to moment of inertia).

In the case of electron, the spin of the electron creates a spin mangetic moment, so

$$\hat{\mu}_s = \frac{1}{2}\frac{e}{m_e}g\hat{s}_z$$

hence the eigenvalues are

$$\mu_s = \frac{e}{2m_e}g\hbar s$$

And thus the total magnetic moment is

$$\vec{\mu}_e = \vec{\mu} + \vec{\mu}_s = \frac{e}{2m_e}(\vec{L}_z + g\vec{s}_z) = \frac{e\hbar}{2m_e}(m_l + gs)\hat{k}$$

We should remember that m_l here refers to the angular momentum without the influence of the spin, therefore it is a whole integer. Also, it should be noted that this means that $\vec{\mu} \not\parallel \vec{J}$, where $\vec{J}$ is the total angular momentum, $\vec{J} = \vec{L}_z + \vec{s}_z$.

Hence the energy of the interaction is

$$E_m = -\vec{\mu}\cdot\vec{B} = -\frac{e\hbar}{2m_e}(m_l + gs)B = -\mu_B(m_l + gs)B$$

Here, constant $\mu_B = \frac{e\hbar}{2m_e}$ is called the Bohr magneton.

The gyromagnetic ration g determination for electron is beyond this course, but it comes out of QED as approximately 2.

3.2.1 Energy Levels in Atom

To further advance on the topic of Zeeman effect, we must consult the energy of different eigenstates of electron in the atom. The main difference in energies of the atom is due to the the main quantum number, n. Bigger n means bigger total energy of the electron. Smaller difference in energy is due to the orbital energy, and therefore due to quantum number l. Finally, even smaller energy difference is due to magnetic moments of the electrons in orbits, which generally correspond to quantum numbers m_l and s.

The orbitals (energy eigenstates of the electron) in the atom are labeled by the main number, latter and a number in the index.

The main number corresponds to the n number. The letter corresponds to the l number in sequence s, p, d, f, ... There l is a whole integer. The index number corresponds to total angular momentum (orbital + spin angular momentum) of the electron in the specified state.

The spin-orbit coupling is an effect due to electron orbiting the nucleus, and therefore does not occur for orbitals with no orbit angular momentum ($l = 0$).

Imagine a frame of reference with stationary electron - the positive nucleus orbits around the electron. This creates a magnetic field. The spin magnetic moment of the electron in this field interacts and different spin has different energy. Usually, this effect is quite weak, and spectral line splitting due to spin-orbit coupling is usually called anomalous Zeeman effect, and occurs without presence of external magnetic field. For example, the sodium doublet is created by spin-orbit splitting of 3p orbital (the transitions to 3s level)
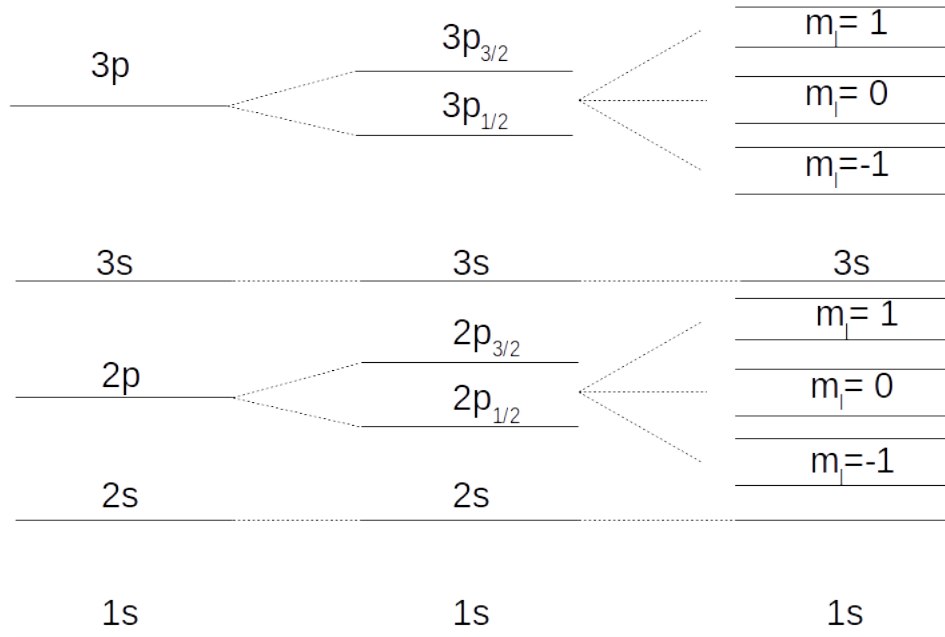
Figure 1: Energy levels in an atom. The levels most left are simply due to n and l. The levels on the right of these are changed by spin-orbit coupling. The most right levels are separated by external magnetic field and different m_l, together with the effect of spin-orbit coupling

3.2.2 Selection Rules

Since the transitions between the energy levels of atoms occurs by absorbtion/fluorescence of a photon, which carries spin angular momentum of $\hbar$, only certain transitions are allowed, due to conservation of angular momentum. Particularly, total angular momentum number j must always change by ± 1, and magnetic angular momentum number m_l must change by ± 1 or stay the same (angular momentum increased for components in xy plane).

3.3 Multiple Electron Atoms

To generally predict behaviour of more than one electron in central Coulombic field is practically impossible. Therefore, models are used to simulate the real situation. Here, we mention the central field approximation.

3.3.1 Central Field Approximation

We can see that as n increases, the distance of the main part of R part of the wavefunction of the electrons from the nucleus increases. Therefore, we can come up with the concept of shells. Shell is a group of energy states of the atom with the same n. They are usually refered to either by numbers or by big latin letters starting from K.
The central field approximation assumes that electrons in the same shell do not interact, and that their wavefunctions behave like wavefunctions of electron around a hydrogen-like atom in given shell, with charge of the nucleus reduced by charge of the electrons in shells with smaller n. Also, central field approximation assumes that Pauli's exclusion principle applies, so that no two electrons exist with all four quantum numbers the same.

3.3.2 Hund's First Rule

We can now imagine filling the orbitals of the atom with electrons. We can imagine that first orbital 1s is filled, after that 2s, then 2p etc. However, there is one exception to this. The spin-orbit coupling interaction is not very strong and in fact more powerful effect is the allignment of electron spins in the same orbital.

Therefore, Hund's first rule states that orbitals are first filled with electrons with the same spin, and only afterwards these start to fill with electrons with opposite spins.

For example, say we have a nucleus with filled 1s and 2s orbitals. Now we add other electrons. First goes to 2p orbital, second also to 2p with the same spin, which means that it has different m_l. Third goes also to 2p with the same spin, hence it has yet a different m_l. Finally, the fourth electron has to have opposite spin and m_l identical with one of the previous electrons.

This procedure can be then used to predict electronic configuration of elements in the periodic table.

4   Free Fermi Gases

Free Fermi gases are systems of fermions within the infinite potential well of finite dimensions. The hamiltonian operator of a wavefunction in this potential well is

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r})$$

where $V(\vec{r}) = 0$ if $\vec{r}$ is position vector of point inside the well and $V(\vec{r}) \to \infty$ if $\vec{r}$ is outside the well. We consider large systems - large potential wells - relative to the wavelengths of the fermions in such systems. This means that, if the medium is isotropic, which is what we usually require, the boundary conditions in parts of the medium must be translatable, and thus periodic - wavefunction of a particle at some end of the part must be the same as at the other end of the part/system.

In order for TISE equation to apply, $\psi(\vec{r}) \to 0$ for $\vec{r}$ outside the system. Therefore, we only need to solve TISE inside the system, where $V = 0$. So, we have

$$\hat{H}\psi = -\frac{\hbar^2}{2m}\nabla^2\psi = E\psi$$

we try to find the eigenfunctions and eigenvalues of this hamiltonian $\hat{H}$. We can notice that

$$[\hat{H}, \hat{p}_i] = \frac{i\hbar^3}{2m}\left(\nabla^2\frac{\partial}{\partial x_i} - \frac{\partial}{\partial x_i}\nabla^2\right) = 0$$

for operator $\hat{p}_i$ of component of momentum in any direction. This means that the hamiltonian operator and momentum operators are compatible, hence they share a common set of eigenfunctions. Eigenfunctions of $\hat{p}_i$ are typically plane waves

$$\psi_p = Ce^{ik_i x_i}$$

where k_i is some wavenumber and C is the normalization constant.

Furthemore, we can notice that

$$\hat{H} = \frac{1}{2m}\left(\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2\right)$$

Which means that there are no other than momentum operators in the hamiltonian. Hence, the eigenfunctions of this free hamiltonian are simply the products of three eigenfunctions of momentum component operators. Some eigenfunction of the hamiltonian is then

$$\psi = Ce^{ik_x x}e^{ik_y y}e^{ik_z z} \tag{39}$$

where C is the normalization constant.

This can be rewritten as

$$\psi = Ce^{i(k_x x + k_y y + k_z z)} = Ce^{i\vec{k}\cdot\vec{r}}$$

where $\vec{k}$ is the wavevector of the fermion.

Now, we invoke the periodic boundary conditions on a cube of the whole system - we require that

$$\psi(x + L_x, y, z) = \psi(x, y, z)$$

$$\psi(x, y + L_y, z) = \psi(x, y, z)$$

$$\psi(x, y, z + L_z) = \psi(x, y, z)$$

These reduce to

$$Ce^{ik_i(x_i + L_i)}e^{i(k_j x_j + k_l x_l)} = Ce^{i(k_i x_i + k_j x_j + k_l x_l)}$$

where $i \neq j \neq l$ and x_i corresponds to some coordinate. This means that

$$e^{ik_i L_i} = 1$$

$$k_i = \frac{2\pi n_i}{L_i}$$

$$k_x = \frac{2\pi n_x}{L_x}, k_y = \frac{2\pi n_y}{L_y}, k_z = \frac{2\pi n_z}{L_z}$$

This means that the eigenstates of the energy are quantized.

Now, if we add fermions to the system, due to exclusion principle, only two fermions can have identical triplet (n_x, n_y, n_z) and these must differ in spin. This means that the more fermions are added to the system, the bigger momenta, and thus energy, they have.

4.1 Fermi Surface

Now, imagine adding the fermions to the system at 0 K. This means that all fermions will be in the state with the lowest energy. However, multiple states correspond to the same energy. To see this, lets find the energy of the fermion in its eigenstate using TISE

$$E\psi = -\frac{\hbar^2}{2m}\nabla^2\Psi = -\frac{\hbar^2}{2m}\nabla^2\left(Ce^{i(k_x x + k_y y + k_z z)}\right) = -\frac{\hbar^2}{2m}\left(-k_x^2 - k_y^2 - k_z^2\right)Ce^{i(k_x x + k_y y + k_z z)} = \frac{\hbar^2|\vec{k}|^2}{2m}\psi$$

$$E = \frac{\hbar^2|\vec{k}|^2}{2m} \tag{40}$$

where $|\vec{k}| = \sqrt{k_x^2 + k_y^2 + k_z^2}$ and $k_i = \frac{2\pi n_i}{L_i}$. We can clearly see that state characterized by $(n_x, n_y, n_z) = (1, 0, 0)$ and another state with $(n_x, n_y, n_z) = (0, 1, 0)$ have the same energy.

If there is a big number of fermions in the system, we can try to further understand how the fermions are distributed between these states.

Consider a map of states in wavevector space. This means a set of points $\vec{k}$ such that each point has coordinates $\vec{k} = \left(\frac{2\pi n_x}{L_x}, \frac{2\pi n_y}{L_y}, \frac{2\pi n_z}{L_z}\right)$. We search for number of states that have energy smaller than some energy E_f. In our map, this corresponds to the number of points inside a sphere of radius $|\vec{k}_f| = \frac{1}{\hbar}\sqrt{2mE_f}$. For a big number of points inside this sphere, this number can be approximated as the volume of the sphere divided by the volume per one point - one state. The volume of the sphere is

$$V_f = \frac{4}{3}\pi|\vec{k}_f|^3 = \frac{4}{3}\pi\left(\frac{2mE_f}{\hbar^2}\right)^{\frac{3}{2}}$$

The volume per one state is the volume of the box created by smallest allowed changes of wavevectors. Labeling it as ΔV

$$\Delta V = \Delta k_x \Delta k_y \Delta k_z = \frac{(2\pi)^3 \Delta n_x \Delta n_y \Delta n_z}{L_x L_y L_z}$$

But, because n_i are simply integers, the smallest change in them is 1. Hence

$$\Delta V = \frac{(2\pi)^3}{V}$$

where V is the volume of our system. This volume is independent of any quantum numbers, which is a important property.

Hence the number of states with energy smaller than E_f is

$$N(E_f) = \frac{V_f}{\Delta V} = \frac{4\pi}{3(2\pi)^3}\left(\frac{2mE_f}{\hbar^2}\right)^{\frac{3}{2}} V$$

But, since each state can be occupied effectively by 2 fermions with opposite spins, the number of fermions in the system with energy smaller than E_f is in fact

$$N(E_f) = 2\frac{V_f}{\Delta V} = \frac{8\pi}{3 \times 8\pi^3}\left(\frac{2mE_f}{\hbar^2}\right)^{\frac{3}{2}} V = \frac{1}{3\pi^2}\left(\frac{2mE_f}{\hbar^2}\right)^{\frac{3}{2}} V \tag{41}$$

or in terms of wavevector

$$N(|\vec{k}_f|) = \frac{1}{3\pi^2} \left(|\vec{k}_f|^2\right)^{\frac{3}{2}} V = \frac{|\vec{k}_f|^3}{3\pi^2} V \tag{42}$$

In case that $N(E_f)$ is the number of all fermions in the system, the corresponding energy E_f is called the Fermi energy of the system. Similarly, $\vec{k}_f$ is called the Fermi wavevector. In our case, the surface of constant energy is surface of constant $|\vec{k}_f|$ because the fermions are free, but this generally does not have to be true. In such case, the surface of constant energy forms a specific surface in the wavevector space, called the Fermi surface.

4.1.1   Lower Dimensions

The same analysis can be carried out for number of states in lower number of dimensions that the fermions move in. Lets start with 2 dimensions.
The analogue of sphere in 2 dimensions is a circle. Therefore, number of states with energy less then E_f is the number of points in the wavevector space that are within circle centered on origin with radius $|\vec{k}_f|$. Hence

$$V_f = \pi |\vec{k}_f|^2 = \pi \frac{2mE_f}{\hbar^2}$$

represents the area of the circle (2D analogue of volume).
The volume (area) per one state is

$$\Delta V = \frac{1}{2}\Delta k_x \Delta k_y = \frac{1}{2}\frac{(2\pi)^2 \Delta n_x \Delta n_y}{L_x L_y} = \frac{2\pi^2}{V}$$

where V is the volume (area) of the system (following the same process of determining square periodic boundary conditions). The factor of $\frac{1}{2}$ accounts for the spin degeneracy (two fermions in one energy state). Hence

$$N(E_f) = \frac{V_f}{\Delta V} = \frac{1}{2\pi}\frac{2mE_f}{\hbar^2} V$$

$$N(|\vec{k}_f|) = \frac{|\vec{k}_f|^2}{2\pi} V$$

Similarly for one dimensional system

$$V_f = 2|k_f| = 2\sqrt{\frac{2mE_f}{\hbar^2}}$$

$$\Delta V = \frac{1}{2}\frac{2\pi\Delta n_x}{L_x} = \frac{\pi}{V}$$

$$N(E_f) = \frac{V_f}{\Delta V} = \frac{2}{\pi}\sqrt{\frac{2mE_f}{\hbar^2}} V$$

$$N(|\vec{k}_f|) = \frac{2|\vec{k}_f|}{\pi} V$$

4.1.2   Density of States

We now the total number of occupied states with energy less then some energy E is

$$N(E) = C_D \left(\frac{2mE}{\hbar^2}\right)^{\frac{D}{2}} V_D$$

where C_D is some dimensionality dependent constant, V_D is the dimensionality dependent volume (or area or length) of the system and D is the number of dimensions.
To find the number of states between energies E and $E + dE$ (labeled as $N'(E)$), we take the difference of total number of states with energy less then $E + dE$ and total number of states with energy less then E. For small dE

$$N'(E) = N(E + dE) - N(E) = \frac{dN}{dE}dE$$

Hence, we can use notation $N'(E) = dN(E)$. We can then see that the total number of states can be reconstructed as

$$N(E) = \int_0^E dN = \int_0^E \frac{dN}{dE} dE$$

We usually mark $\frac{dN}{dE} = g(E)$ and call $g(E)$ the density of states. In 3D

$$g(E) = \frac{d}{dE}\left(C_D \left(\frac{2mE}{\hbar^2}\right)^{\frac{D}{2}}\right) V_D = C_D \frac{D}{2}\left(\frac{2m}{\hbar^2}\right)^{\frac{D}{2}} E^{\frac{D-2}{2}} V_D = \frac{D}{2}\frac{N(E)}{E} \tag{43}$$

To express the density of states in terms of wavevector, we must be a bit careful. We want the density of states to behave as

$$dN(|\vec{k}|) = g(|\vec{k}|)d|\vec{k}|$$

labeling $|\vec{k}| = k$

$$g(k) = \frac{dN}{dk} = \frac{dE}{dk}\frac{dN}{dE} = \frac{d}{dk}\left(\frac{\hbar^2 k^2}{2m}\right)\frac{D}{2}N(k)\frac{2m}{\hbar^2 k^2} = D\frac{N(k)}{k} = DC_D k^{D-1} V_D$$

Hence in 3D

$$g(E) = \frac{1}{3\pi^2}\frac{3}{2}\left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}} E^{\frac{1}{2}} V_3 = \frac{1}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}}\sqrt{E} V_3$$

$$g(k) = 3\frac{1}{3\pi^2}k^2 V_3 = \frac{k^2}{\pi^2} V_3$$

In 2D

$$g(E) = \frac{1}{2\pi}\left(\frac{2m}{\hbar^2}\right) V_2$$

$$g(k) = 2\frac{1}{2\pi}k V_2 = \frac{k}{\pi} V_2$$

And in 1D

$$g(E) = \frac{1}{\pi}\sqrt{\frac{2m}{\hbar^2}}\frac{1}{\sqrt{E}} V_1$$

$$g(k) = \frac{2}{\pi} V_1 = \frac{2V_1}{\pi}$$

4.1.3   Total Energy

Total energy of the system can be found by summing together product of energy E and number of states at this energy $dN(E)$ for all energies E occupied in the system. This is equal to

$$E_T = \int_0^{E_f} E dN = \int_0^{E_f} E\frac{dN}{dE} dE = \int_0^{E_f} E g(E) dE \tag{44}$$

We can evaluate this as

$$E_T = \int_0^{E_f} E C_D \frac{D}{2}\left(\frac{2m}{\hbar^2}\right)^{\frac{D}{2}} E^{\frac{D-2}{2}} V_D dE = C_D \frac{D}{2}\left(\frac{2m}{\hbar^2}\right)^{\frac{D}{2}} V_D \int_0^{E_f} E^{\frac{D}{2}} dE =$$

$$= C_D \frac{D}{2}\left(\frac{2m}{\hbar^2}\right)^{\frac{D}{2}} V_D \frac{2}{D+2} E_f^{\frac{D+2}{2}} = \frac{D}{D+2} N(E_f) E_f$$

Hence the average energy

$$< E >= \frac{E_T}{N(E_f)} = \frac{D}{D+2} E_f$$

4.1.4 Fermi Energy and Wavevector

To find the explicit magnitude of the Fermi wavevector and Fermi energy, we use similar integral as before. We know

$$N(E_f) = N_T = \int_0^{E_f} dN = \int_0^{E_f} g(E)dE$$

where N_T is the total number of fermions in the system. This means

$$N_T = \int_0^{E_f} C_D \frac{D}{2} \left(\frac{2m}{\hbar^2}\right)^{\frac{D}{2}} E^{\frac{D-2}{2}} V_D dE = C_D \frac{D}{2} \left(\frac{2m}{\hbar^2}\right)^{\frac{D}{2}} V_D \int_0^{E_f} E^{\frac{D-2}{2}} dE =$$

$$= C_D \frac{D}{2} \left(\frac{2m}{\hbar^2}\right)^{\frac{D}{2}} V_D \frac{2}{D} E_f^{\frac{D}{2}} = N(E_f)$$

which was expected. We can turn the equation around to find

$$E_f^{\frac{D}{2}} = \frac{1}{C_D} \left(\frac{\hbar^2}{2m}\right)^{\frac{D}{2}} \frac{N_T}{V_D} = \left(\frac{\hbar^2}{2m}\right)^{\frac{D}{2}} \frac{n}{C_D}$$

where n is the number density of fermions in the system. Hence

$$E_f = \frac{\hbar^2}{2m} \left(\frac{n}{C_D}\right)^{\frac{2}{D}} = \frac{\hbar^2}{2m} \left(\frac{\rho}{mC_D}\right)^{\frac{2}{D}}$$

where ρ is the density of the fermions. In terms of Fermi wavenumber

$$k_f^2 = \left(\frac{n}{C_D}\right)^{\frac{2}{D}}$$

$$k_f = \left(\frac{n}{C_D}\right)^{\frac{1}{D}} = \left(\frac{\rho}{mC_D}\right)^{\frac{1}{D}}$$

4.2 Free Electron Gases in Metals

One classical example of free Fermi gases are free electron gases in metals. Electrons in metals are perfectly dissociated from individual atoms and free to move in the crystal lattice of the metal. The nuclei of the atoms only form a constant positive potential background and electrons are thus effectively trapped in a potential well. Therefore, they obey relationships derived for free Fermi gases.

The Fermi energy for electrons in metals is generally quite high. Usually, the number density of electrons is around $n \approx 10^{28}$ m^{-3}, which leads to $E_f \approx 1.7eV$, which corresponds to thermal oscillations around 20 000 K.

Some specifics for electrons in metals are now derived.

4.2.1 Conduction

We are interested in a cause behind conduction and its mathematical realisation in our free Fermi gases formalism. To advance, consider TDSE for free electrons

$$i\hbar \frac{\partial}{\partial t} \Psi = -\frac{\hbar^2}{2m} \nabla^2 \Psi$$

Taking the complex conjugate

$$-i\hbar \frac{\partial}{\partial t} \Psi^* = -\frac{\hbar^2}{2m} \nabla^2 \Psi^*$$

Multiplying the first equation by Ψ^* and the second by Ψ, we obtain set of equations

$$i\hbar \Psi^* \frac{\partial}{\partial t} \Psi = -\frac{\hbar^2}{2m} \Psi^* \nabla^2 \Psi$$

$$-i\hbar \Psi \frac{\partial}{\partial t} \Psi^* = -\frac{\hbar^2}{2m} \Psi \nabla^2 \Psi^*$$

Subtracting the second equation from the first one

$$i\hbar \left(\Psi^* \frac{\partial}{\partial t} \Psi + \Psi \frac{\partial}{\partial t} \Psi^* \right) = -\frac{\hbar^2}{2m} \left(\Psi^* \nabla^2 \Psi - \Psi \nabla^2 \Psi^* \right)$$

From vector analysis, we now that

$$\nabla \cdot (f\vec{v}) = (\nabla f) \cdot \vec{v} + f(\nabla \cdot \vec{v})$$

Hence we can say that

$$\nabla \cdot (\Psi^* \nabla \Psi) = (\nabla \Psi^*) \cdot (\nabla \Psi) + \Psi^* \nabla^2 \Psi$$

Similarly

$$\nabla \cdot (\Psi \nabla \Psi^*) = (\nabla \Psi) \cdot (\nabla \Psi^*) + \Psi \nabla^2 \Psi^*$$

Subtracting the first equation from the second one, we arrive at

$$\nabla \cdot (\Psi^* \nabla \Psi) - \nabla \cdot (\Psi \nabla \Psi^*) = \Psi^* \nabla^2 \Psi + (\nabla \Psi^*) \cdot (\nabla \Psi) - \Psi \nabla^2 \Psi^* - (\nabla \Psi) \cdot (\nabla \Psi^*)$$

Since the scalar product is commutative and divergence operator is distributive with respect to addition

$$\nabla \cdot (\Psi^* \nabla \Psi - \Psi \nabla \Psi^*) = \Psi^* \nabla^2 \Psi - \Psi \nabla^2 \Psi^*$$

Hence the original equation becomes

$$i\hbar \frac{\partial}{\partial t} (\Psi^* \Psi) = -\frac{\hbar^2}{2m} \nabla \cdot (\Psi^* \nabla \Psi - \Psi \nabla \Psi^*)$$

$$\frac{\partial}{\partial t} (\Psi^* \Psi) = -\nabla \cdot \left(\frac{\hbar}{2mi} (\Psi^* \nabla \Psi - \Psi \nabla \Psi^*) \right)$$

This resembles a continuity equation $\frac{\partial \rho}{\partial t} = -\nabla \cdot \vec{j}$ where $\vec{j} = \frac{\hbar}{2mi} (\Psi^* \nabla \Psi - \Psi \nabla \Psi^*)$ is some flux density. It turns out that this flux density in fact corresponds to the transport of the particle.
In our case, where $\Psi = Ce^{i\vec{k} \cdot \vec{r}}$

$$\vec{j} = \frac{\hbar}{2mi} \left(C^* e^{-i\vec{k} \cdot \vec{r}} \nabla Ce^{i\vec{k} \cdot \vec{r}} - Ce^{i\vec{k} \cdot \vec{r}} \nabla C^* e^{-i\vec{k} \cdot \vec{r}} \right) =$$

$$= CC^* \frac{\hbar}{2mi} \left(e^{-i\vec{k} \cdot \vec{r}} (ik_x \hat{i} + ik_y \hat{j} + ik_z \hat{k}) e^{i\vec{k} \cdot \vec{r}} - e^{i\vec{k} \cdot \vec{r}} (-ik_x \hat{i} - ik_y \hat{j} - ik_z \hat{k}) e^{-i\vec{k} \cdot \vec{r}} \right) =$$

$$= |C|^2 \frac{\hbar}{2m} (\vec{k} + \vec{k}) = |C|^2 \frac{\hbar}{m} \vec{k}$$

Hence the transport of the electrons is dependent on their wavevector. But if the wavevectors for different electrons are symmetrically spaced around the origin, as we have assumed so far, there is no net transport. This means that in a conducting free electron metal the centre of the spherical shell of constant energy is effectively shifted in some direction. The transport than occurs in the direction of the shift, and is proportional to the magnitude of the shift.
To show this more explicitly, we need to find the normalization constant for given plane waves with

$$\vec{k} = \frac{2\pi n_x}{L_x} \hat{i} + \frac{2\pi n_y}{L_y} \hat{j} + \frac{2\pi n_z}{L_z} \hat{k}$$

The normalization condition is

$$1 = \iiint_V \Psi^* \Psi dV$$

$$1 = \int_0^{L_x} \int_0^{L_y} \int_0^{L_z} C^* e^{-i\vec{k} \cdot \vec{r}} Ce^{i\vec{k} \cdot \vec{r}} dz dy dx = |C|^2 \int_0^{L_x} \int_0^{L_y} \int_0^{L_z} dz dy dx = |C|^2 L_x L_y L_z = |C|^2 V$$

$$|C|^2 = \frac{1}{V}$$

This has simplest form for $C = \frac{1}{\sqrt{V}}$. We see that the normalization constant is independent of $\vec{k}$, which will be critical in the next step.

To find the overall transport flux density $\vec{J}$ of the system of free electrons, we need to sum the flux density for transport of individual electrons for all electrons. That is

$$\vec{J} = \sum_i \vec{j}_i = \sum_i \frac{\hbar}{2mi}(\Psi_i^* \nabla \Psi_i - \Psi_i \nabla \Psi_i^*) = \frac{\hbar}{m}\sum_i |C_i|^2 \vec{k}_i = \frac{\hbar}{Vm}\sum_i \vec{k}_i$$

Since the points in wavevector space are spaced symmetrically around the displacement $\vec{k}_d$ for conducting medium

$$\sum_i \vec{k}_i = \vec{k}_d$$

Hence

$$\vec{J} = \frac{\hbar}{Vm}\vec{k}_d$$

And we could speculate that the current density $\vec{j}_q$ (current per unit area) goes as

$$\vec{j}_q = e\vec{J} = \frac{\hbar e}{Vm}\vec{k}_d$$

4.2.2 Heat Capacity

The average occupacy of the energy states in with energy E at temperature $T \neq 0$ for fermions is called the Fermi-Dirac distribution and has form

$$\eta(E,T) = \frac{1}{e^{\frac{(E-E_f)}{k_B T}} + 1} \tag{45}$$

where k_B is the Boltzmann constant and E_f is the Fermi energy.

At $T \to 0$ limit, this distribution becomes a step function with $\eta(E,0) \to 1$ for $E < E_f$ and $\eta(E,0) \to 0$ for $E > E_f$.

Using this distribution, the integral over all states becomes

$$N_T = \int_0^\infty \eta(E,T)g(E)dE$$

where $g(E)$ is the density of states. We can see that at $T \to 0$, this reduces to already derived

$$N_T = \int_0^\infty \eta(E,0)g(E)dE = \int_0^{E_f} g(E)dE$$

Now consider the case of non-zero temperature. First, we find the value of $\eta(E_f, T)$

$$\eta(E_f,T) = \frac{1}{e^{\frac{E_f - E_f}{k_B T}} + 1} = \frac{1}{2}$$

Now, consider a point $E = E_f + \Delta E$. Fermi-Dirac distribution is

$$\eta(E_f + \Delta E, T) = \frac{1}{e^{\frac{E_f + \Delta E - E_f}{k_B T}} + 1} = \frac{1}{e^{\frac{\Delta E}{k_B T}} + 1} = \frac{e^{\frac{-\Delta E}{k_B T}}}{1 + e^{\frac{-\Delta E}{k_B T}}} = 1 - \frac{1}{e^{\frac{-\Delta E}{k_B T}} + 1} = 1 - \eta(E_f - \Delta E, T)$$

$$\eta(E_f + \Delta E, T) = 1 - \eta(E_f - \Delta E, T)$$

This means that the distribution is somewhat symmetrical around E_f.

Because $E_f >> k_B T$, the distribution tends to 1 as $E \to 0$ and to 0 as $E \to \infty$. Consider a point $E = E_f + k_B T$

$$\eta(E_f + k_B T, T) = \frac{1}{e^1 + 1} \approx 0.27$$

Hence the distributions main feature for common temperatures will be in rather small interval of approximately $[E_f - k_B T, E_f + k_B T]$. Otherwise, the distribution is pretty much the same as for $T = 0$.

The total energy of the system at temperature T is then

$$E_T(T) = \int_0^\infty E\eta(E,T)g(E)dE = \int_0^{E_f} E\eta(E,T)g(E)dE + \int_{E_f}^\infty E\eta(E,T)g(E)dE$$

The energy at $T = 0$ is

$$E_T(0) = \int_0^\infty E\eta(E,0)g(E)dE = \int_0^{E_f} Eg(E)dE$$

Hence the difference in the energy at $T = 0$ and at $T = T$ is

$$E_T(T) - E_T(0) = \int_0^{E_f} Eg(E)(\eta(E,T) - 1)dE + \int_{E_f}^\infty E\eta(E,T)g(E)dE$$

Since $\eta(E,T) - 1$ for $E < E_f$ is non-zero only for a small interval before E_f, and $\eta(E,T)$ for $E > E_f$ is non-zero only for a small interval after E_f, the density of states can be assumed to be constant along this interval, i. e. $g(E) \approx g(E_f)$ Hence

$$E_T(T) - E_T(0) = \int_0^{E_f} Eg(E)(\eta(E,T) - 1)dE + \int_{E_f}^\infty E\eta(E,T)g(E)dE \approx$$

$$\approx g(E_f)\left[\int_0^{E_f} E(\eta(E,T) - 1)dE + \int_{E_f}^\infty E\eta(E,T)dE\right]$$

Performing a substitution $E = E_f - E'$ for the first integral and $E = E_f + E'$ for the second integral, we obtain.

$$E_T(T) - E_T(0) \approx g(E_f)\left[\int_{E_f}^0 (E_f - E')(\eta(E_f - E',T)(-dE')) + \int_0^\infty (E_f + E')\eta(E_f + E',T)dE'\right] =$$

$$= g(E_f)\left[\int_0^{E_f} (E' - E_f)(1 - \eta(E_f - E',T))dE' + \int_0^\infty (E' + E_f)\eta(E_f + E',T)dE'\right] =$$

Using the symmetry of Fermi-Dirac distribution and linearity of integration

$$= g(E_f)\left[\int_0^{E_f} E'\eta(E_f + E',T)dE' - E_f\int_0^{E_f} \eta(E_f + E',T)dE' + \int_0^\infty E'\eta(E_f + E',T)dE' +\right.$$

$$\left. + E_f\int_0^{E_f} \eta(E_f + E',T)dE' + E_f\int_{E_f}^\infty \eta(E_f + E',T)dE'\right]$$

Now, since E_f is much bigger than k_BT, we can assume that $\eta(E_f + E',T)$ for any $E' \geq E_f$ is practically zero (this is equal to $\eta(2E_f,T)$). The last integral therefore goes to zero, and as the second one and fourth subtract each other out, we are left with

$$E_T(T) - E_T(0) = g(E_f)\left[\int_0^{E_f} E'\eta(E_f + E',T)dE' + \int_0^\infty E'\eta(E_f + E',T)dE'\right]$$

Again, since $\eta(E_f + E',T)$ goes to zero as an exponential, we can assume that $E'\eta(E_f + E',T)$ will go to zero as well for $E' \geq E_f$. Hence we can add

$$0 = g(E_f)\int_{E_f}^\infty E'\eta(E_f + E',T)dE'$$

to the right hand side of the equation, and we obtain

$$E_T(T) - E_T(0) = g(E_f)\left[\int_0^{E_f} E'\eta(E_f + E',T)dE' + \int_{E_f}^\infty E'\eta(E_f + E',T)dE' + \int_0^\infty E'\eta(E_f + E',T)dE'\right]$$

$$E_T(T) - E_T(0) = 2g(E_f)\int_0^\infty E'\eta(E_f + E',T)dE'$$

To proceed further, we consider the explicit form of this integral

$$\int_0^\infty E'\eta(E_f + E',T)dE' = \int_0^\infty E'\frac{1}{1 + e^{\frac{E_f + E' - E_f}{k_BT}}}dE' = \int_0^\infty E'\frac{1}{1 + e^{\frac{E'}{k_BT}}}dE'$$

Substituting $x = \frac{E'}{k_B T}$

$$\int_0^\infty E' \frac{1}{1 + e^{\frac{E'}{k_B T}}} dE' = k_B^2 T^2 \int_0^\infty \frac{x}{1 + e^x} dx$$

The integral is a standadr integral and converges to $\frac{\pi^2}{12}$. Hence

$$E_T(T) - E_T(0) = 2g(E_f)k_B^2 T^2 \frac{\pi^2}{12} = \frac{\pi^2}{6} g(E_f)k_B^2 T^2$$

Hence the heat capacity of the electron gas is

$$C = \lim_{\Delta T \to 0} \frac{E_T(T + \Delta T) - E_T(T)}{\Delta T} = \lim_{\Delta T \to 0} \frac{E_T(T + \Delta T) - E_T(0) - (E_T(T) - E_T(0))}{\Delta T} =$$

$$= \frac{\partial (E_T(T) - E_T(0))}{\partial T} = \frac{\pi^2}{3} g(E_f)k_B^2 T$$

Substituting in for $g(E_f)$ (in 3D)

$$g(E_f) = \frac{1}{2\pi^2} \left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}} \sqrt{E_f} V$$

where

$$E_f = \frac{\hbar^2}{2m} \left(\frac{3\pi^2 N}{V}\right)^{\frac{2}{3}}$$

where N is the number of fermions in the system.
Hence

$$g(E_f) = \frac{1}{2\pi^2} \left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}} \left(\frac{\hbar^2}{2m}\right)^{\frac{1}{2}} \left(\frac{3\pi^2 N}{V}\right)^{\frac{1}{3}} V = \frac{1}{2\pi^2} \frac{2m}{\hbar^2} \left(3\pi^2 N\right)^{\frac{1}{3}} V^{\frac{2}{3}}$$

And thus

$$C = \frac{\pi^2}{3} \frac{1}{2\pi^2} \frac{2m}{\hbar^2} \left(3\pi^2 N\right)^{\frac{1}{3}} V^{\frac{2}{3}} k_B^2 T = \frac{1}{3} \frac{m}{\hbar^2} \left(3\pi^2 N\right)^{\frac{1}{3}} V^{\frac{2}{3}} k_B^2 T$$

4.2.3   Magnetic Susceptibility

When magnetic field is applied to the metal, the electrons with opposite spins do not have the same energy anymore. Instead, energy of some electrons goes up by $\gamma = \mu_B B$, where $\mu_B = \frac{e\hbar}{2m_e}$ is the Bohr magneton and B is the magnitude of magnetic induction field applied, and energy of other electrons goes down by γ
We can express the density of states as sum of density of states of particles with one direction of spin that has increased energy and the density of state of other particles with decreased energy

$$g(E) = \frac{dN}{dE} = \frac{d(N_+ + N_-)}{dE} = \frac{dN_+}{dE} + \frac{dN_-}{dE} = g_+(E) + g_-(E)$$

Since the energy of the interaction γ is usually small, the number of states that gets elevated by γ over the original level is approximately

$$dN_+ \approx \left.\frac{dN_+}{dE}\right|_{E_f} \gamma = g_+(E_f)\gamma$$

Similarly, the number of states that decrease in energy compared to situation without external field is approximately

$$dN_- \approx -\left.\frac{dN_-}{dE}\right|_{E_f} (-\gamma) = g_-(E_f)\gamma$$

As a consequence of this shift in energies, some of the particles that had higher energy will switch spins in order to decrease total energy. Therefore, all dN_+ particles will switch to opposite spins. Therefore, there is now $2dN_+$ particles more in the opposite spin (dN_+ added with opposite spin and dN_+ lost with original spin). Hence, the overall magnetisation is

$$M = \mu_B 2dN_+ \approx \mu_B 2g_+(E_f)\gamma$$

44

Since before the application of B field there was no magnetisation, we had $g_+(E_f) = g_-(E_f) = \frac{1}{2}g(E_f)$. Therefore

$$M = \mu_B g(E_f)\mu_B B = g(E_f)\mu_B^2 B \tag{46}$$

Hence the (total) magnetic susceptibility is

$$\chi = \frac{\partial M}{\partial B} = g(E_f)\mu_B^2 \tag{47}$$

This corresponds to linear magnetisation with no hysterisis, which corresponds to our assumption that the electrons do not interact.

4.3 Compact Stars

Stars in certain period near the end of their life undergo contraction to a point when only the pressure due to fermions being to close holds the star from becoming a black hole.

Where does this pressure come from? Modeling the fermions as free (which is usually quite good approximation) and star as homogeneous (only somewhat true during the end of life of the star), the energy of one fermion is

$$E = \frac{\hbar^2}{2m}(k_x^2 + k_y^2 + k_z^2) = \frac{h^2}{2m}\left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2}\right)$$

we can then see that if we decrease any of L_x, L_y or L_z, the energy of the fermion increases. This energy has to be supplied somehow, and thus the compression of the system must be done with some force. This is equivalent effect to a pressure from the system opposing the compression.

Now, imagine that the gravity of the star compresses it and the fermion pressure, called also the degeneracy pressure, opposes this compression. To find whether there can exist a stable equilibrium when these effects are equal, we may consider the gravitational potential energy of the star and kinetic energy of the fermions. The gravitational potential energy is (assuming homogeneous star)

$$U = \int_0^M -\frac{GM'(r)dm}{r}$$

where $M'(r)$ is the mass of the part of the star with distance from the centre smaller than r, dm is the mass of the spherical shell of thickness dr at radius r, M is the total mass of the star and G is the Newton's gravitational constant.

For homogeneous star

$$M'(r) = \frac{4}{3}\pi r^3 \rho$$

where ρ is density of the star, and

$$dm = 4\pi r^2 \rho dr$$

Therefore

$$U = \int_0^R -G\frac{16}{3}\pi^2\rho^2 r^4 dr = -G\frac{16}{15}\pi^2\rho^2 R^5 = -\frac{3}{5}\frac{G}{R}\left(\frac{4}{3}\pi R^3\rho\right)^2 = -\frac{3}{5}\frac{GM^2}{R}$$

The kinetic energy of the fermions is (in 3D)

$$T = \frac{3}{5}NE_f$$

where N is the number of fermions in the system. Using the average mass of the star per fermion (total mass divided by number of fermions holding the star) m_{pf} and relationship for Fermi energy

$$T = \frac{3}{5}\frac{M}{m_{pf}}\frac{\hbar^2}{2m}\left(3\pi^2\frac{N}{\frac{4}{3}\pi R^3}\right)^{\frac{2}{3}} = \frac{3}{5}\frac{M}{m_{pf}}\frac{\hbar^2}{2m}\left(\frac{9\pi M}{4m_{pf}}\right)^{\frac{2}{3}}\frac{1}{R^2} = \frac{3}{5}\left(\frac{9\pi}{4}\right)^{\frac{2}{3}}\left(\frac{M}{m_{pf}}\right)^{\frac{5}{3}}\frac{\hbar^2}{2mR^2}$$

where m is the mass of the fermion, which is not necceseraly equal to m_{pf}, because the star might not be made just from the fermions.

Therefore, both energies are simply functions of mass and radius of the star. To search for the minimum of total energy with respect to radius R (for a fixed mass - mass has to be conserved for the star - which is also an assumption), we write the total energy as

$$H = T + U = \frac{\alpha}{R^2} - \frac{\beta}{R}$$

$$\frac{\partial H}{\partial R} = -2\frac{\alpha}{R^3} + \frac{\beta}{R^2} = 0$$

$$\beta R = 2\alpha$$

$$R = \frac{2\alpha}{\beta}$$

where $\alpha = \frac{3}{5}\left(\frac{9\pi}{4}\right)^{\frac{2}{3}}\left(\frac{M}{m_{pf}}\right)^{\frac{2}{3}}\frac{\hbar^2}{2m}$ and $\beta = \frac{3}{5}GM^2$. Hence

$$R = \frac{2}{GM^2}\left(\frac{9\pi}{4}\right)^{\frac{2}{3}}\left(\frac{M}{m_{pf}}\right)^{\frac{5}{3}}\frac{\hbar^2}{2m} = \left(\frac{9\pi}{4}\right)^{\frac{2}{3}}\frac{\hbar^2}{mGM^{\frac{1}{3}}m_{pf}^{\frac{5}{3}}}$$

For white dwarfs, which are stars about the with mass about tenth the mass of the sun, the fermions supporting the stars are electrons. Assuming that there is approximately one proton for each electron in the star (star was originally from hydrogen), $m_{pf} \approx m_p$ for white dwarfs, as electrons are very light. Using the values stated

$$R \approx 5 \times 10^7 \text{ m}$$

which is comparable to the size of the Earth (about ten times as big). For more massive white dwarfs, the size of the star starts to be similar to the earths size.

However, as the star gets smaller and smaller, the electrons get faster and faster, until they reach a high relativistic limit. In this limit, the energy of the electron is approximated as

$$E \approx \hbar k c$$

Hence, the total energy of the fermions is

$$T = \frac{3}{5}E_f \approx \frac{3}{5}\hbar c k_f = \frac{3}{5}\hbar c\left(3\pi^2\frac{3M}{m_{pf}4\pi^2 R^3}\right)^{\frac{1}{3}} = \frac{3}{5}\hbar c\left(\frac{9\pi}{4}\right)^{\frac{1}{3}}\left(\frac{M}{m_{pf}}\right)^{\frac{1}{3}}\frac{1}{R}$$

This means that the total energy is

$$H = \frac{\alpha'}{R} - \frac{\beta}{R} = \frac{\alpha' - \beta}{R}$$

For previous values, $\alpha' \approx 1.8 \times 10^{-7}$ J m and $\beta \approx 1.6 \times 10^{48}$ J m, which means that $\alpha' - \beta$ is negative, and therefore the star can decrease the energy as long as it keeps decreasing its radius. This can lead to creation to the neutron star, where the star is supported by neutrons instead of electrons. In this case, electrons react with the protons to form neutrons, and hence $m_{pf} \approx m = m_n \approx m_p$. Using the non-relativistic limit again (for slow neutrons) and parameters as before

$$R \approx 3 \times 10^4 \text{ m}$$

which is about the size of a city.

4.4   Nuclei

The key observation here is that the density of nuclei is close to constant. This means that the volume of the nucleus is proportional to the atomic number A, which is the sum of the number of protons and neutrons, Z and N, respectively.

The nucleus is held together by the strong force, which has only very small range - effectively only directly adjacent particles can be binded by it. Therefore, the binding energy of the strong force goes as $a_1 A - a_2 A^{\frac{2}{3}}$ since the particles on the surface ($A^{\frac{2}{3}}$)do not have the same amount of adjacent particles as the particles in the middle. For each other particle, there is some constant binding energy term.

Other part of the total energy of the nucleus is created by the Coulomb electrostatic interaction, which goes as $-a_3\frac{Z(Z-1)}{A^{\frac{1}{3}}}$, as only the protons interact.

Other important part is created by the proton and neutron movement. This is the part that can actually be modelled by the free Fermi gas model.

Since protons and neutrons are different fermions, they both have different Fermi energies, E_{FZ} and E_{FN}. The free fermion model predicts that the total kinetic energy will be

$$E_T = \frac{3}{5}ZE_{FZ} + \frac{3}{5}NE_{FN}$$

Substituting in for Fermi energy using the number density of protons and neutrons

$$E_T = \frac{3}{5}\left(Z\frac{\hbar^2}{2m_Z}\left(3\pi^2 n_Z\right)^{\frac{2}{3}} + N\frac{\hbar^2}{2m_N}\left(3\pi^2 n_N\right)^{\frac{2}{3}}\right) \approx \frac{3}{5}\frac{\hbar^2}{2m_u}(3\pi^2)^{\frac{2}{3}}\left(Zn_Z^{\frac{2}{3}} + Nn_N^{\frac{2}{3}}\right)$$

Where we used that masses of proton and neutron are quite close and both are close to atomic mass unit $m_u \approx 1.67 \times 10^{-27}$ kg.
Using

$$n_Z = \frac{Z}{V}$$

and similar expression for n_N, we obtain

$$E_T = \frac{3}{5}\frac{\hbar^2}{2m_u}\left(\frac{3\pi^2}{V}\right)^{\frac{2}{3}}\left(Z^{\frac{5}{3}} + N^{\frac{5}{3}}\right)$$

This can be rewritten as

$$E_T = \frac{3}{5}\frac{\hbar^2}{2m_u}\left(\frac{3\pi^2}{V}\right)^{\frac{2}{3}}\left(\left(\frac{Z+N+Z-N}{2}\right)^{\frac{5}{3}} + \left(\frac{Z+N-(Z-N)}{2}\right)^{\frac{5}{3}}\right)$$

Since $Z + N = A$

$$E_T = \frac{3}{5}\frac{\hbar^2}{2m_u}\left(\frac{3\pi^2}{V}\right)^{\frac{2}{3}}\left(\frac{A}{2}\right)^{\frac{5}{3}}\left(\left(1 + \frac{Z-N}{A}\right)^{\frac{5}{3}} + \left(1 - \frac{Z-N}{A}\right)^{\frac{5}{3}}\right)$$

Usually, $A >> Z - N$ (for big nuclei). Using binomial expansion

$$\left(1 + \frac{(Z-N)}{A}\right)^{\frac{5}{3}} \approx 1 + \frac{5}{3}\frac{(Z-N)}{A} + \frac{\frac{5}{3}\times\frac{2}{3}}{2}\frac{(Z-N)^2}{A^2} = 1 + \frac{5}{3}\frac{(Z-N)}{A} + \frac{5}{9}\frac{(Z-N)^2}{A^2}$$

Similarly

$$\left(1 - \frac{(Z-N)}{A}\right)^{\frac{5}{3}} \approx 1 - \frac{5}{3}\frac{(Z-N)}{A} + \frac{5}{9}\frac{(Z-N)^2}{A^2}$$

Hence

$$E_T = \frac{3}{5}\frac{\hbar^2}{2m_u}\left(\frac{3\pi^2}{V}\right)^{\frac{2}{3}}\left(\frac{A}{2}\right)^{\frac{5}{3}}\left(2 + 2\times\frac{5}{9}\frac{(Z-N)^2}{A^2}\right) =$$

$$= \frac{3}{5}\frac{\hbar^2}{2m_u}\left(\frac{3\pi^2\frac{A}{2}}{V}\right)^{\frac{2}{3}}\left(A + \frac{5}{9}\frac{(Z-N)^2}{A}\right)$$

We can see that the factor in front of the last term in brackets is the Fermi energy for case when there is exactly the same number of neutrons and protons. This is why the correction in brackets is called the asymmetry term. Using $\rho = \frac{M}{V} \approx \frac{Am_u}{V}$, where M is the total mass of the nucleus, we obtain

$$E_T = \frac{3}{5}\frac{\hbar^2}{2m_u}\left(3\pi^2\frac{\rho}{2m_u}\right)^{\frac{2}{3}}\left(A + \frac{5}{9}\frac{(Z-N)^2}{A}\right)$$

Since we observe that ρ is approximately constant for all nuclei, we can now determine $\bar{\epsilon}_F$ as a constant and therefore have

$$E_T = \frac{3}{5}\bar{\epsilon}_F A + \frac{3}{5}\bar{\epsilon}_F\frac{(Z-N)^2}{A}$$

Usually, the first term is subtracted from the bonding energy of the strong force, and is included in first term $a_1 A$. The second term is a term on its own. Due to effective mass and other considerations, we generalize it to $a_4\frac{(N-Z)^2}{A}$.
The last term in the energy of the nucleus is so called pairing term. This term represents the energy due to pairing of neutrons and protons. Right now it is more of a phenomenological term, as it for some reason decreases as $A^{-\frac{4}{3}}$. The full term is

$$a_5\frac{\left[(-1)^N + (-1)^Z\right]}{2A^{\frac{4}{3}}}$$

This term is added to the binding energy, hence if there is even number of protons and neutrons, the binding energy is increased (both types of particles in pairs), while for both even, the binding energy is decreased.

The full total binding energy of the nucleus is then

$$U_B = a_1 A - a_2 A^{\frac{2}{3}} - a_3 \frac{Z(Z-1)}{A^{\frac{1}{3}}} - a_4 \frac{(N-Z)^2}{A} + a_5 \frac{(-1)^N + (-1)^Z}{2A^{\frac{4}{3}}}$$

4.5   Limitations of the Free Fermion Model

Free fermion model is very useful because it is very easy to understand and use. However, it is rarely satisfied perfectly. In fact, the white dwarfs are the best real objects similar to free fermion system. In metals, the behaviour is basically similar, but other processes take place as well. This is usually represented in form of effective mass of electron - the electron is not taken as having mass $m = m_e \approx 511$ keV but rather at some mass m^*, which depends on the material. Otherwise, the model is usually quite precise for well conducting metals.

5   Crystals and Lattices

5.1   Bravais Lattices

Many real materials form crystals. In description of crystals, the term of Bravais lattice makes occurance. Bravais lattice is any set of objects ordered in space in such a way that the environment around each object is identical for all objects in the lattice.

We usually say that the points at which the objects are placed are the lattice points, and the objects are the lattice basis.

Usually, the requirement of identical environment leads to periodical structure.

We usually try to express real lattices as Bravais lattices using collection of atoms/ions as the basis of the lattice (and sometimes the collection consists of a single atom/ion).

Another term usually mentioned in crystalography is the unit and primitive unit cells. Primitive unit cell is the smallest collection of ions that can be the basis of a Bravais lattice. Unit cell is any collection that can be the basis of a Bravais lattice, which can be for example a combination of primitive unit cells.

Surprisingly, there is not many possible Bravais lattices - 5 in 2D and 17 in 3D. We will not list them explicitly here, but some basic, such as simple cubic, body-centeret cubic (BCC), face-centered cubic (FCC) and hexagonal close packed (HCP) are hopefully known well. It should be noted that for the identical particles, both BCC and FCC can be interpreted as simple cubic with multiatom basis.

5.2   Reciprocal Lattice Vectors

Since the crystal structure is usually periodical, all fields due to the atoms must have a periodical structure as well. This means that we can express any function of position $\vec{r}$ as Fourier sum

$$f(\vec{r}) = \sum_{\vec{G}} f_{\vec{G}} e^{i\vec{G}\cdot\vec{r}}$$

where the sum goes over all possible values of $\vec{G}$.

Now assume that vectors $\vec{a}_1$, $\vec{a}_2$ and $\vec{a}_3$ form vector basis of the lattice - any translation from a lattice point by vector $\vec{a} = m\vec{a}_1 + n\vec{a}_2 + o\vec{a}_3$ will only translate to a different lattice point, thus not creating any change. This means that

$$f(\vec{r} + \vec{a}) = f(\vec{r})$$

This means that

$$e^{i\vec{G}\cdot(\vec{r}+\vec{a})} = e^{i\vec{G}\cdot\vec{r}}$$

$$e^{i\vec{G}\cdot\vec{a}} = 1$$

$$\vec{G} \cdot \vec{a} = 2\pi N$$

where N is an integer.

We could then express $\vec{G}$ in basis of $\vec{a}_1$, $\vec{a}_2$ and $\vec{a}_3$ and search for vectors that satisfy the above condition. But, what is usually done is that we define a new basis that has the same direction of vectors but different

magnitudes. This is the reciprocal lattice vector basis and the vectors can be derived from the original basis vectors as follows (in 3D)

$$\vec{b}_1 = 2\pi \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}$$

$$\vec{b}_2 = 2\pi \frac{\vec{a}_3 \times \vec{a}_1}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}$$

$$\vec{b}_3 = 2\pi \frac{\vec{a}_1 \times \vec{a}_2}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}$$

The important property of these vectors is that they have the directions of original basis. This can be shown for each vector explicitly

$$\vec{b}_1 \cdot \vec{a}_1 = \frac{2\pi}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} (\vec{a}_2 \times \vec{a}_3) \cdot \vec{a}_1 = 2\pi$$

Denoting $\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3) = V$ as the volume of the unit cell (which it indeed is), we have for other two vectors

$$\vec{b}_1 \cdot \vec{a}_2 = \frac{2\pi}{V} (\vec{a}_2 \times \vec{a}_3) \cdot \vec{a}_2$$

Since $\vec{a}_2 \times \vec{a}_3 \perp \vec{a}_2$ (property of vector product)

$$\vec{b}_1 \cdot \vec{a}_2 = 0$$

Similarly

$$\vec{b}_1 \cdot \vec{a}_3 = \frac{2\pi}{V} (\vec{a}_2 \times \vec{a}_3) \cdot \vec{a}_3 = 0$$

And similarly for other reciprocal lattice vectors

$$\vec{b}_2 \cdot \vec{a}_1 = \frac{2\pi}{V} (\vec{a}_3 \times \vec{a}_1) \cdot \vec{a}_1 = 0$$

$$\vec{b}_2 \cdot \vec{a}_2 = \frac{2\pi}{V} (\vec{a}_3 \times \vec{a}_1) \cdot \vec{a}_2 = \frac{2\pi}{V} \vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3) = 2\pi$$

where I used a cyclic permutation.

$$\vec{b}_2 \cdot \vec{a}_3 = \frac{2\pi}{V} (\vec{a}_3 \times \vec{a}_1) \cdot \vec{a}_3 = 0$$

$$\vec{b}_3 \cdot \vec{a}_1 = \frac{2\pi}{V} (\vec{a}_1 \times \vec{a}_2) \cdot \vec{a}_1 = 0$$

$$\vec{b}_3 \cdot \vec{a}_2 = \frac{2\pi}{V} (\vec{a}_1 \times \vec{a}_2) \cdot \vec{a}_2 = 0$$

$$\vec{b}_3 \cdot \vec{a}_3 = \frac{2\pi}{V} (\vec{a}_1 \times \vec{a}_2) \cdot \vec{a}_3 = \frac{2\pi}{V} \vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3) = 2\pi$$

Therefore, we can summarize our findings as

$$\vec{b}_m \cdot \vec{a}_n = 2\pi \delta_{mn}$$

where δ_{mn} is the Kronecker delta. Using this definition, we can find the value of the triple product $\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)$. We start with the inner vector product

$$\vec{b}_2 \times \vec{b}_3 = \frac{(2\pi)^2}{V^2} (\vec{a}_3 \times \vec{a}_1) \times (\vec{a}_1 \times \vec{a}_2) = \frac{(2\pi)^2}{V^2} ([(\vec{a}_3 \times \vec{a}_1) \cdot \vec{a}_2]\vec{a}_1 - [(\vec{a}_3 \times \vec{a}_1) \cdot \vec{a}_1]\vec{a}_2) =$$

$$= \frac{(2\pi)^2}{V^2} [\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)]\vec{a}_1 = \frac{(2\pi)^2}{V} \vec{a}_1$$

and

$$\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3) = \frac{(2\pi)^2}{V} \vec{b}_1 \cdot \vec{a}_1 = \frac{(2\pi)^3}{V}$$

which is the volume of the reciprocal lattice unit cell.

If we look at this and previous equation, we can see that

$$\vec{a}_1 = \frac{V}{(2\pi)^2}(\vec{b}_2 \times \vec{b}_3) = 2\pi \frac{\vec{b}_2 \times \vec{b}_3}{\frac{(2\pi)^3}{V}} = 2\pi \frac{\vec{b}_2 \times \vec{b}_3}{\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)}$$

Similarly for other vectors

$$\vec{b}_3 \times \vec{b}_1 = \frac{(2\pi)^2}{V^2}(\vec{a}_1 \times \vec{a}_2) \times (\vec{a}_2 \times \vec{a}_3) = \frac{(2\pi)^2}{V^2}(((\vec{a}_1 \times \vec{a}_2) \cdot \vec{a}_3)\vec{a}_2 - ((\vec{a}_1 \times \vec{a}_2))\vec{a}_2)\vec{a}_3 =$$

$$= \frac{(2\pi)^2}{V}\vec{a}_2$$

$$\vec{a}_2 = 2\pi \frac{\vec{b}_3 \times \vec{b}_1}{\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)}$$

$$\vec{b}_1 \times \vec{b}_2 = \frac{(2\pi)^2}{V^2}(\vec{a}_2 \times \vec{a}_3) \times (\vec{a}_3 \times \vec{a}_1) = \frac{(2\pi)^2}{V^2}(((\vec{a}_2 \times \vec{a}_3) \cdot \vec{a}_1)\vec{a}_3 - ((\vec{a}_2 \times \vec{a}_3) \cdot \vec{a}_3)\vec{a}_1) =$$

$$= \frac{(2\pi)^2}{V}\vec{a}_3$$

$$\vec{a}_3 = 2\pi \frac{\vec{b}_1 \times \vec{b}_2}{\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)}$$

Hence the inverse transforms are exactly identical to the forward transforms, hence why we call the lattice reciprocal - applying the transform twice gives back the original lattice and the vectors have dimensions that are reciprocal dimensions of the original lattice.

To end our thought, all $\vec{G}$ vectors must be then sum of integer multiples of the reciprocal lattice vectors, as (for $\vec{G} = h\vec{b}_1 + k\vec{b}_2 + l\vec{b}_3$)

$$\vec{G} \cdot \vec{a} = \sum_{i=1}^{3} h\vec{b}_1 \cdot \vec{a}_i + k\vec{b}_2 \cdot \vec{a}_{i+1} + l\vec{b}_3 \cdot \vec{a}_{i+2}$$

where cyclic permutation $a_4 \rightarrow a_1$ occurs. Hence

$$\vec{G} \cdot \vec{a} = 2\pi(h + k + l)$$

which is an integer multiple of 2π, which is what we originally required.

5.3   Bragg Diffraction and Miller Indices

Integers h, k and l are called the miller indices. Each vector in the reciprocal lattice can be then defined using only these three indices, hence $\vec{G} = \vec{G}_{hkl}$. Also, such a vector defines a plane in the reciprocal lattice. This plane is also a plane in the real lattice, as basis vectors of reciprocal lattice and real lattice are respectively parallel.

When X-Rays (or other waves) scatter from these planes, Bragg diffraction occurs. The Bragg diffraction has maxima at

$$m\lambda = 2\sin\theta d_{hkl}$$

where m is an integer, λ is wavelength of the incident wave, θ is the angle of incidence and d_{hkl} is the distance between the planes defined by $\vec{G}_{hkl}$ which have overlapping atoms when looking in direction of $\vec{G}_{hkl}$.

The distance must therefore also satisfy

$$nd_{hkl} = \hat{G}_{hkl} \cdot \vec{r}$$

where n is an integer, $\hat{G}_{hkl}$ is the unit reciprocal lattice vector and $\vec{r}$ is the real lattice vector.

5.4 Nearly Free Electron Model

The crystal structure is the principal idea behind a nearly free electron model for electrons in solids. It assumes that the variation of potential due to stationary positive ions in the lattice is relatively small. This means that the potential can be described by a function

$$V(\vec{r}) = \sum_{h,k,l} V_{hkl} e^{i\vec{G}_{hkl} \cdot \vec{r}}$$

Now, for clarity, consider a 1D case. Here

$$V = V(x) = \sum_h V_h e^{iG_h x} = \sum_h V_h e^{i\frac{\pi h}{L} x}$$

where $2L$ is the length of the unit cell from $-L$ to L in 1D.

If the variation in V is small, it usually means that coefficients V_h are small as well. Since the series converges, only few first coefficients for small h. We can try to construct a primitive model where only the first two coefficent are important, and therefore

$$V \approx V_1 e^{i\frac{2\pi}{a} x} + V_{-1} e^{-i\frac{2\pi}{a} x} = V_1 e^{iG_1 x} + V_{-1} e^{iG_1 x}$$

Now we further assume that $V_{-1} = V_1 = V$, which basically just requires that the function V is even, which makes sense if it is centered on an atom at $x = 0$ (for any other function, the x direction would not be isotropic).

We can try to find approximate eigenstates of the hamiltonian operator by using non-eigenstates which we would expect to find. Since the perturbation of the free hamiltonian by V is small, we can try to use plane waves to simulate this. To make the model symplest possible, we only allow two specific states with wavenumbers k and $k - G_1$. We use the matrix modelling to find corresponding eigenvalues of energy. We therefore need to determine the elements of the hamiltonian matrix.

$$< \frac{e^{ikx}}{\sqrt{2L}} |\hat{H}| \frac{e^{ikx}}{\sqrt{2L}} > = \frac{1}{2L} \int_{-L}^{L} e^{-ikx} \left(-\frac{\hbar^2}{2m^*} \frac{\partial^2}{\partial x^2} + V e^{iG_1 x} + V e^{-iG_1 x} \right) e^{ikx} dx =$$

$$= \frac{1}{2L} \int_{-L}^{L} e^{-ikx} \left(\frac{\hbar^2 k^2}{2m^*} e^{ikx} + V e^{ix(k+G_1)} + V e^{ix(k-G_1)} \right) dx$$

Because both k and G_1 correspond to wavenumbers for waves periodic on $-L$ to L in x, we have orthogonality relations

$$\int_{-L}^{L} e^{-ikx} dx = 2L \delta_{k0}$$

where δ_{k0} is the Kronecker delta.

Hence

$$< \frac{e^{ikx}}{\sqrt{2L}} |\hat{H}| \frac{e^{ikx}}{\sqrt{2L}} > = \frac{\hbar^2 k^2}{2m^*}$$

Similarly

$$< \frac{e^{i(k-G_1)x}}{\sqrt{2L}} |\hat{H}| \frac{e^{i(k-G_1)x}}{\sqrt{2L}} > = \frac{1}{2L} \int_{-L}^{L} e^{-i(k-G_1)x} \left(\frac{\hbar^2}{2m^*} \frac{\partial^2}{\partial x^2} + V e^{iG_1 x} + V e^{-iG_1 x} \right) e^{i(k-G_1)x} dx =$$

$$= \frac{1}{2L} \int_{-L}^{L} e^{-i(k-G_1)x} \left(\frac{\hbar^2 (k-G_1)^2}{2m^*} e^{i(k-G_1)x} + V e^{ikx} + V e^{i(k-2G_1 x)} \right) dx = \frac{\hbar^2 (k-G_1)^2}{2m^*}$$

Now, for the off-diagonal elements.

$$< \frac{e^{ikx}}{\sqrt{2L}} |\hat{H}| \frac{e^{i(k-G_1)x}}{\sqrt{2L}} > = \frac{1}{2L} \int_{-L}^{L} e^{-ikx} \left(\frac{\hbar^2 (k-G_1)^2}{2m^*} e^{i(k-G_1)x} + V e^{ikx} + V e^{i(k-2G_1)x} \right) dx = V$$

$$< \frac{e^{i(k-G_1)x}}{\sqrt{2L}} |\hat{H}| \frac{e^{ikx}}{\sqrt{2L}} > = \frac{1}{2L} \int_{-L}^{L} e^{-i(k-G_1)x} \left(\frac{\hbar^2 k^2}{2m^*} e^{ikx} + V e^{i(k+G_1)x} + V e^{ikx} \right) dx = V$$

Hence the matrix is

$$\mathbf{H} = \begin{pmatrix} \frac{\hbar^2 k^2}{2m^*} & V \\ V & \frac{\hbar^2 (k-G_1)^2}{2m^*} \end{pmatrix}$$

Lets mark $E_k = \frac{\hbar^2 k^2}{2m^*}$ and consistently with this also $E_{k-G_1} = \frac{\hbar^2(k-G_1)^2}{2m^*}$. To find the eigenvalues

$$\begin{vmatrix} E_k - \lambda & V \\ V & E_{k-G_1} - \lambda \end{vmatrix} = 0$$

$$(E_k - \lambda)(E_{k-G_1} - \lambda) - V^2 = 0$$

$$\lambda^2 - (E_k + E_{k-G_1})\lambda + E_k E_{k-G_1} - V^2 = 0$$

$$\lambda = \frac{E_k + E_{k-G_1}}{2} \pm \frac{1}{2}\sqrt{(E_k + E_{k-G_1})^2 - 4(E_k E_{k-G_1} - V^2)}$$

$$\lambda = \frac{E_k + E_{k-G_1}}{2} \pm \sqrt{\left(\frac{E_k - E_{k-G_1}}{2}\right) + V^2}$$

Thus for states with similar E_k and E_{k-G_1} (states around $k = \frac{G_1}{2}$), the difference of energies of the two states is $2V$, which is purely material dependent and finite. This means that the electrons cannot transfer between these two states, and hence cannot change direction without requiring energy.

But, conduction in materials occurs by swaping the direction of the momentum but not changing the energy of the electrons much. In this case, some wavenumbers cannot change the momentum without the change in energy - they must climb certain energy gap.

Proper treatment leads to the formalism of bands and gaps.

5.4.1   Bands and Gaps

The electrons in solids that have some periodic disruptive potential acting on the electrons stimulate the creation of electron bands. These bands are intervals of allowed wavenumbers which have small separation in energy (by $\Delta E = \frac{\hbar^2 \pi^2}{2m^* L^2}$). Between these bands, the band gaps are present, where electrons with given wavenumbers and energies do not exist.

Further high in energy, electrons with some wavenumbers and higher energies occur again.

When conduction occurs, it slightly shifts the wavevectors in some direction, while slightly increasing the energy. But for materials that have big band gap and filled band with lower energy, the electrons would have to overcome the band gap energy barrier to conduct, which very hard.

Effectively, the electric potential then pushes against the degeneracy pressure of the electrons, which is very strong.

6   Semiconductors

Semi-conductors are materials which have filled or nearly filled bottom electron band. This means that only very few electrons that overcome the band gap due to thermal excitations can conduct.

The number of these conductive electrons is given by Boltzmann distribution as (assuming only the electrons close to the band gap starting energy)

$$N_e = N_0 e^{-\frac{\Delta}{k_B T}}$$

where Δ is the width of the energy band gap and N_0 is some number of electrons close to the energy gap. We do not really need this number, and therefore we will just write that

$$N_e \propto e^{-\frac{\Delta}{k_B T}}$$

However, this electron that gets excited to the higher band (usually called the conduction band) leaves behind effectively a free spot (free wavevector value) which can be filled by an electron of higher energy. This empty spot is often called a hole and in fact behaves as a quasi-particle. Imagine that we plot the electron energies on a vertical axis. Then, while the electron tries to flow down the axis to minimise its energy, the hole needs to flow up to minimize the energy of the other electrons which sequentially take its place.

This hole therefore always drifts to the top of the lower band (usually called the valence band). The hole can conduct in this conduction band, and thus the total number of conducting particles is

$$N = N_e + N_h = 2N_e = 2N_0 e^{-\frac{\Delta}{k_B T}}$$

6.1 Doping

We can artificially add other elements to the semiconductor to change its conductive properties. For a semiconductor with four valence electrons (Si or Ge), we could add either Al (3 valence electrons) or As/P (five valence electrons). These atoms get implemented in the crystal structure of the semiconductor, and thus either have one stable electron extra (for 3 valence electrons elements) or one less (for five valence electrons). Also, they create a permanent hole/electron that can conduct and is not caused by thermal excitation.

Consider now the case of 3 valency element. The extra electron that is bound to the element creates effective fixed negative charge, and therefore the energy of this fixed electron is slightly higher than the energy of other electrons from the semiconductor. Therefore, the Fermi energy of the semiconductor is shifted up by this electron, sitting small amount of energy above the valence band.

Analogously, the positive static 5 valency element still attracts the electron a little bit (they form so called exciton pair). This means that the electron connected to this element has slightly lower energy than the other electrons from the semiconductor. Since the electron from this atom goes directly to the conduction band, the Fermi energy of such doped semiconductor lies a small bit under the bottom of the conduction band.

Summarizing, we have two possibilities for doping. Either we dope with 5 valency elements, putting extra electrons into the system, which can conduct, while increasing the Fermi energy just under the conduction band, or dope with 3 valency elements, putting extra holes into the system and increasing the Fermi energy slightly above the valence band. The first type of doping is the n-doping, the second is p-doping (by charge of the doped charge carrier/ also called the majority carrier).

6.2 Aligning Materials

When the materials are aligned, the stable state is achieved when their Fermi energies are aligned. This might cause some initial electron flux, but the stable state is alignment of the Fermi energies.

Therefore, if we put together one p-doped (p-type) and n-doped (n-type) semiconductor, the bands align in such a way that the bottom of the conduction band of the n-type is almost aligned with the top of the valence band of the p-type (difference due to interaction energies of excitons - static positive/negative charge).

In such settings, the electron that gets to the conduction band of the p-type material quickly travels to the top of the conduction band of the n-type material, as the bottom of the conduction band in the n-type is now much lower. Similarly, any hole present in the n-type quickly travels to the p-type, as it can bubble up through the electrons.

The number of electrons that get excited to the conduction band of the p-type per time is

$$n_{e0} \propto e^{-\frac{\Delta}{k_B T}}$$

similarly, the number of holes created by excitation of the electrons in the n-type material is

$$n_{h0} \propto e^{-\frac{\Delta}{k_B T}}$$

In a steady state solution, these numbers are aligned, so

$$n_{e0} = n_{h0}$$

or, by multiplying by charge e and taking the current in direction from n-type to p-type, we have

$$i_{e0} = i_{h0}$$

Hence the total current (of positive charge) flowing from n-type to p-type in the thermal equilibrium is

$$I_0 = i_{h0} + i_{e0} = 2i_{h0}$$

This current is also called the generation current.

This would imply that there is always some non-zero current, which does not occur in reality, because there is one more type of current that can occur. The electrons from the n-type that sit in the conduction band have non-zero probability of climbing the potential to conduction band in the p-type. The potential they climb is approximately Δ, so the current will be again

$$i_{er} \propto -e^{-\frac{\Delta}{k_B T}}$$

where minus sign stands for the electrons travelling in $n \to p$ direction and having negative charge. Similarly, holes can climb the potential down to the n-type with the same probability. Hence

$$I_r = -i_{hr} - i_{er} = -2i_{hr}$$

In thermal equilibrium, there is no net current, so we must have

$$I_r + I_0 = 0$$

or

$$i_{hr} = i_{h0}$$

and

$$i_{er} = i_{e0}$$

The I_r current is also called the recombination current, as the electrons/ holes arriving to the opposite type material quickly recombine and anihilate.

6.3   Conduction

Now, imagine biasing the n-type with voltage V. This shifts the energy of electrons in the n-type by eV up, thus the electrons/holes from the recombination current must only climb potential difference of $\Delta - eV$ now, so

$$I_r = I_{r,eq} e^{-\frac{(-eV)}{k_B T}} = I_{r,eq} e^{\frac{eV}{k_B T}}$$

where $I_{r,eq}$ is the recombination current in equilibrium.
The generation current remains the same as in equilibrium

$$I_g = I_{g,eq} = -I_{r,eq}$$

Hence the total current across the junction is

$$I = I_g + I_r = I_{r,eq} \left(e^{\frac{eV}{k_B T}} - 1 \right)$$

which is the familiar behaviour of a diode.
The last remark should be made on recombination. The recombination is followed by decrease in energy of the electron, which needs to be radiated away. In order to be radiated away (by a photon), the transition must be allowed by selection rules (conservation of angular momentum). As it turns out, in silicon, the selection rules forbid most recombination transitions crossing the band gap. This caused problems as the pile up of electrons stopped the current from flowing. It then took longer for people to pursue this research, hence why we discovered semiconductor devices so late.

7   Standard Model

7.1   Klein-Gordon Equation

The quantum mechanics as we derived it so far has one major problem - it is not special relativistic theory. To incorporate a special relativity into the theory, we need to use principle of correspondence to find some relativistic quantum equation and some principle from non-relativistic quantum mechanics to be consistent with this new equation in order for the new theory to be consistent with the old theory.
Motivations to discover this theory are mainly particle physics and explanations of some foreign concepts, such as spin or antiparticles.
Historically, the Klein-Gordon equation was the first try to do this. It can be derived from the relativistic expression for total energy

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

Replacing the variables with corresponding operators (and mass with just multiplication by mass)

$$-\hbar^2 \frac{\partial^2}{\partial t^2} \Psi = -\hbar^2 c^2 \nabla^2 \Psi + m^2 c^4 \Psi$$

We can try to scale the units so that all variables are in in terms of energy, i.e.

$$t' = \frac{t}{\hbar}, x' = \frac{x}{\hbar c}, m' = mc^2$$

These are called the natural units and I will drop the prime in the following section. The Klein-Gordon equation in natural units then becomes

$$\frac{\partial^2}{\partial t^2}\Psi - \nabla^2\Psi + m^2\Psi = 0$$

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2 + m^2\right)\Psi = 0 \tag{48}$$

This is the Klein-Gordon equation. To make connection to the non-relativistic quantum mechanics, we would like some form of probability density to be derivable from this equation. The probability density should also satisfy the continuity equation

$$\frac{\partial\rho}{\partial t} = -\nabla \cdot \vec{j}$$

where $\vec{j}$ is some flux density and ρ is the probability density. We found this equation originally by taking a complex conjugate of TDSE and comparing to the original, so we can try to do the same thing here. Taking complex conjugate

$$\frac{\partial^2}{\partial t^2}\Psi^* - \nabla^2\Psi^* + m^2\Psi^2 = 0$$

Now, we multiply the new equation by Ψ and the original equation by Ψ^*, and thus obtain

$$\Psi\frac{\partial^2}{\partial t^2}\Psi^* - \Psi\nabla^2\Psi^* + \Psi m^2\Psi^* = 0 \tag{49}$$

$$\Psi^*\frac{\partial^2}{\partial t^2}\Psi - \Psi^*\nabla^2\Psi + \Psi^* m^2\Psi = 0 \tag{50}$$

As before, we can subtract (49) from (50) to get

$$\Psi^*\frac{\partial^2}{\partial t^2}\Psi - \Psi\frac{\partial^2}{\partial t^2}\Psi^* - \Psi^*\nabla^2\Psi + \Psi\nabla^2\Psi^* = 0$$

Consider that

$$\frac{\partial}{\partial t}\left(\Psi^*\frac{\partial}{\partial t}\Psi - \Psi\frac{\partial}{\partial t}\Psi^*\right) = \frac{\partial}{\partial t}\Psi^*\frac{\partial}{\partial t}\Psi + \Psi^*\frac{\partial^2}{\partial t^2}\Psi - \frac{\partial}{\partial t}\Psi\frac{\partial}{\partial t}\Psi^* - \Psi\frac{\partial^2}{\partial t^2}\Psi^* = \Psi^*\frac{\partial^2}{\partial t^2}\Psi - \Psi\frac{\partial^2}{\partial t^2}\Psi^*$$

and

$$\nabla \cdot (\Psi^*\nabla\Psi - \Psi\nabla\Psi^*) = \nabla\Psi^* \cdot \nabla\Psi + \Psi^*\nabla^2\Psi - \nabla\Psi \cdot \nabla\Psi^* - \Psi\nabla^2\Psi^* = \Psi^*\nabla^2\Psi - \Psi\nabla^2\Psi^*$$

Hence we get

$$\frac{\partial}{\partial t}\left(\Psi^*\frac{\partial}{\partial t}\Psi - \Psi\frac{\partial}{\partial t}\Psi^*\right) = -\nabla \cdot (\Psi^*\nabla\Psi - \Psi\nabla\Psi^*)$$

which is our continuity equation with $\rho = \Psi^*\frac{\partial}{\partial t}\Psi - \Psi\frac{\partial}{\partial t}\Psi^*$ and $\vec{j} = \Psi^*\nabla\Psi - \Psi\nabla\Psi^*$. However, this cannot be interpreted as probability density equation, as ρ can very well be negative, which is non-sensical for probability density.

Therefore, Klein-Gordon equation is probably not the equation we are looking for. However, lets try one more example - try to calculate the energy of plane waves.

The plane wave will have a form like (using $\omega t = \hbar\omega\frac{t}{\hbar} = Et'$)

$$\Psi = e^{i(\vec{k}\cdot\vec{r} - Et)}$$

Pluging this into the Klein-Gordon equation

$$-E^2\Psi + |\vec{k}|^2\Psi + m^2\Psi = 0$$

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

But, there is no clear reason why E should be a positive definite, which is again a problem. Also, the total energy is not in any way dependent on the spin of the particle, which we would like it to be (at least to provide degenerate solutions for free particle). Therefore, we need a different equation

7.2   Dirac Equation

Dirac equation tries to be equivalent with Klein-Gordon equation but going one step deeper. Dirac equation states that

$$i\hbar\frac{\partial}{\partial t}\Psi = (\vec{\alpha}\cdot\vec{p}+\beta m)\Psi \tag{51}$$

Or, in natural units and substituting for momentum operator

$$i\frac{\partial}{\partial t}\Psi = (-i\vec{\alpha}\cdot\nabla + \beta m)\Psi$$

Hence

$$\left(i\frac{\partial}{\partial t} + i\vec{\alpha}\cdot\nabla - \beta m\right)\Psi = 0$$

Consider now operating with $i\frac{\partial}{\partial t} - i\vec{\alpha}\cdot\nabla + \beta m$ on the equation. We get

$$\left(i\frac{\partial}{\partial t} - i\vec{\alpha}\cdot\nabla + \beta m\right)\left(i\frac{\partial}{\partial t} + i\vec{\alpha}\cdot\nabla - \beta m\right)\Psi = 0$$

$$\left(-\frac{\partial^2}{\partial t^2} - \frac{\partial}{\partial t}\vec{\alpha}\cdot\nabla - i\frac{\partial}{\partial t}\beta m + \vec{\alpha}\cdot\nabla\frac{\partial}{\partial t} + (\vec{\alpha}\cdot\nabla)^2 + i\vec{\alpha}\cdot\nabla\beta m + i\beta m\frac{\partial}{\partial t} + i\beta m\vec{\alpha}\cdot\nabla - \beta^2 m^2\right)\Psi = 0$$

Assuming that neither of the operators depends explicitly on time, we derive

$$\left(-\frac{\partial^2}{\partial t^2} + (\vec{\alpha}\cdot\nabla)^2 + im(\vec{\alpha}\cdot\nabla\beta + \beta\vec{\alpha}\cdot\nabla) - \beta^2 m^2\right)\Psi = 0$$

Multiplying by -1

$$\left(\frac{\partial^2}{\partial t^2} - (\vec{\alpha}\cdot\nabla)^2 - im(\vec{\alpha}\cdot\nabla\beta + \beta\vec{\alpha}\cdot\nabla) + \beta^2 m^2\right)\Psi = 0$$

This starts to look like Klein-Gordon equation. To make it look exactly, we need to require

$$\beta^2 = 1$$

$$(\vec{\alpha}\cdot\nabla)^2 = \nabla^2$$

and

$$\vec{\alpha}\cdot\nabla\beta + \beta\vec{\alpha}\cdot\nabla = 0$$

We can see that the last operation is hardly possible without the spatial dependence of β, which would be very weird, since β should be associated with mass only. Dirac solved this problem by assuming that neither β nor components of $\vec{\alpha}$ are numbers, but rather that these are matrices and Ψ is a vector. Then, the conditions are

$$\beta^2 = I$$

where I is the identity matrix

$$(\vec{\alpha}\cdot\nabla)^2 = \left(\alpha_1\frac{\partial}{\partial x} + \alpha_2\frac{\partial}{\partial y} + \alpha_3\frac{\partial}{\partial z}\right)^2 = \nabla^2$$

Which is equivalent to

$$\alpha_1^2\frac{\partial^2}{\partial x^2} + \alpha_2^2\frac{\partial^2}{\partial y^2} + \alpha_3^2\frac{\partial^2}{\partial z^2} + (\alpha_1\alpha_2 + \alpha_2\alpha_1)\frac{\partial^2}{\partial x\partial y} + (\alpha_1\alpha_3 + \alpha_3\alpha_1)\frac{\partial^2}{\partial x\partial z} + (\alpha_2\alpha_3 + \alpha_3\alpha_2)\frac{\partial^2}{\partial y\partial z} = \nabla^2 =$$

$$= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

Which leads to set of conditions

$$\alpha_i^2 = I \tag{52}$$

where $i \in \{1, 2, 3\}$ and

$$\alpha_1\alpha_2 + \alpha_2\alpha_1 = \alpha_1\alpha_3 + \alpha_3\alpha_1 = \alpha_2\alpha_3 + \alpha_3\alpha_2 = 0 \tag{53}$$

Therefore, pairs of components of $\vec{\alpha}$ need to be anti-commutating. From the last condition (assuming no spatial dependence of β)

$$\vec{\alpha} \cdot \nabla \beta + \beta \vec{\alpha} \cdot \nabla = 0$$

$$\alpha_1 \beta \frac{\partial}{\partial x} + \alpha_2 \beta \frac{\partial}{\partial y} + \alpha_3 \beta \frac{\partial}{\partial z} + \beta \alpha_1 \frac{\partial}{\partial x} + \beta \alpha_2 \frac{\partial}{\partial y} + \beta \alpha_3 \frac{\partial}{\partial z} = 0$$

$$(\alpha_1 \beta + \beta \alpha_1) \frac{\partial}{\partial x} + (\alpha_2 \beta + \beta \alpha_2) \frac{\partial}{\partial y} + (\alpha_3 \beta + \beta \alpha_3) \frac{\partial}{\partial z} = 0$$

Which again requires

$$\alpha_i \beta + \beta \alpha_i = 0 \tag{54}$$

Matrices obeying the numbered conditions were known and they are called the Pauli matrices.

7.3 Pauli Matrices

Pauli matrices are the following 2x2 matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

To check whether (52) is satisfied for these matrices, consider general matrix equation

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

The square of general 2x2 matrix is

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^2 = \begin{pmatrix} a^2 + bc & b(a + d) \\ c(a + d) & d^2 + bc \end{pmatrix}$$

Hence we have four equations

$$a^2 + bc = 1$$

$$d^2 + bc = 1$$

$$b(a + d) = 0$$

$$c(a + d) = 0$$

From the last two equations, we have two possible situations. Either $a + d \neq 0$ and then $b = c = 0$, or $a + d = 0$. In the first situation, the other two equations become

$$a^2 = d^2 = 1$$

And hence we have our first two solutions as

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

with the first one identical to σ_3 and the second one is trivial identity matrix solution. Now for the case when $a + d = 0$. This means that $a = -d$ and $a^2 = d^2$, and therefore we are left with one equation

$$a^2 + bc = 1$$

$$a = \pm\sqrt{1 - bc}$$

In order to keep things simple, we can search only for solutions where all coeffecents are whole numbered. One such solution is $b = c = \pm 1$, which leads to $a = d = 0$ and hence solution

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

which is σ_1. Similarly, we could choose $b = -i$ and $c = i$ so we would receive still $a = d = 0$ and solution

$$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

which is σ_2. Therefore we see that Pauli matrices are some of the roots of equation $A^2 = I$, where A is the unknown 2x2 matrix.

Other property we need to check is whether these matrices are anti-commutative. To check this, we carry out the calculations explicitly

$$\sigma_1\sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

$$\sigma_2\sigma_1 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$$

Hence

$$\sigma_1\sigma_2 + \sigma_2\sigma_1 = 0$$

Similarly

$$\sigma_1\sigma_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

$$\sigma_3\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

$$\sigma_1\sigma_3 + \sigma_3\sigma_1 = 0$$

$$\sigma_2\sigma_3 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

$$\sigma_3\sigma_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}$$

$$\sigma_2\sigma_3 + \sigma_3\sigma_2 = 0$$

Hence the Pauli matrices satisfy (53). But, now, we need to find a matrix β such that it is anti-commutative with each of the Pauli matrices. This turns out to be very hard to do in 2x2 matrices. But, we can use the property of block multiplication of matrices to avoid this problem and solve for β and α_i in 4x4 matrices. The block multiplication states that for matrices A,B,C,D and A',B',C',D', the following relation applies

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} = \begin{pmatrix} AA' + BC' & AB' + BD' \\ CA' + CC' & CB' + DD' \end{pmatrix}$$

Hence the block multiplication of matrices can be carried out in the same way as normal matrix multiplication, but taking care of the order of operations, as matrices are not necesserily commutative. Now consider a matric product such as

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} aI & bI \\ cI & dI \end{pmatrix}$$

where A,B,C and D are matrices and I is the identity matrix and a,b,c,d are scalars. Since scalar multiplication is commutative with respect to matrix multiplication and I commutes with any matrix by definition, this block multiplication is carried out in exactly the same way as if the matrices A, B, C and D were just some numbers and I was one.

Following this train of thought, we can construct the matrix β as one of the Pauli matrices but instead of numbers defined above we will have these numbers multiplying identity matrix, so for example

$$\beta_1 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$$

or

$$\beta_2 = \begin{pmatrix} 0 & -iI \\ iI & 0 \end{pmatrix}$$

We then construct matrices α so that they are Pauli block matrix created from original Pauli matrices, so for example

$$\alpha_{11} = \begin{pmatrix} 0 & \sigma_1 \\ \sigma_1 & 0 \end{pmatrix}$$

58

$$\alpha_{12} = \begin{pmatrix} 0 & \sigma_2 \\ \sigma_2 & 0 \end{pmatrix}$$

or

$$\alpha_{33} = \begin{pmatrix} \sigma_3 & 0 \\ 0 & -\sigma_3 \end{pmatrix}$$

Then, each element of a set $A_j = \{k \in \{1,2,3\} : \alpha_{jk}\}$ where $j \in \{1,2,3\}$ is guaranteed to be anti-commutative with β_i, $i \neq j$, since the Pauli matrices are anti-commutative. Usually, the set with $i = 1$ and $j = 3$ is chosen, only the sign of the α matrices is switched, so we have

$$\beta = \beta_1 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$$

$$\alpha_i = -\alpha_{3i} = \begin{pmatrix} -\sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix}$$

which is guaranteed to satisfy (54). Strictly speaking, we need to check whether the α_i matrices now satisfy (52) and (53).

$$\alpha_i^2 = \begin{pmatrix} -\sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} \begin{pmatrix} -\sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} = \begin{pmatrix} \sigma_i^2 & 0 \\ 0 & \sigma_i^2 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} = I_4$$

where I_4 is the 4x4 identity matrix. Now for $i \neq j$

$$\alpha_i \alpha_j = \begin{pmatrix} -\sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} \begin{pmatrix} -\sigma_j & 0 \\ 0 & \sigma_j \end{pmatrix} = \begin{pmatrix} \sigma_i \sigma_j & 0 \\ 0 & \sigma_i \sigma_j \end{pmatrix}$$

and similarly

$$\alpha_j \alpha_i = \begin{pmatrix} \sigma_j \sigma_i & 0 \\ 0 & \sigma_j \sigma_i \end{pmatrix}$$

and hence

$$\alpha_i \alpha_j + \alpha_j \alpha_i = \begin{pmatrix} \sigma_i \sigma_j + \sigma_j \sigma_i & 0 \\ 0 & \sigma_i \sigma_j + \sigma_j \sigma_i \end{pmatrix} = 0$$

Therefore, we have found a set of 4x4 matrices that satisfies the given conditions on α_i and β. We could also chosen different template Pauli matrix for β and α_i and we would receive the same results, but the one mentioned above was used in lectures.

Before we move away from Pauli matrices, we quickly calculate the eigenvalues and eigenvectors of the Pauli matrices.

7.3.1 Eigenvectors and Eigenvalues of Pauli matrices

To find the eigenvalues of σ_1

$$\begin{vmatrix} -\lambda_1 & 1 \\ 1 & -\lambda_1 \end{vmatrix} = 0$$

$$\lambda_1^2 - 1 = 0$$

$$\lambda_1 = \pm 1$$

So, the eigenvectors satisfy

$$\begin{pmatrix} \mp 1 & 1 \\ 1 & \mp 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\mp a + b = 0$$

$$b = \pm a$$

And so the normalized eigenvectors with corresponding eigenvalues $\lambda = \pm 1$ are

$$\hat{e}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}$$

Similarly for σ_2

$$\begin{vmatrix} -\lambda_2 & -i \\ i & -\lambda_2 \end{vmatrix} = 0$$

$$\lambda_2^2 - 1 = 0$$

$$\lambda_2 = \pm 1$$

which are the same eigenvalues as for σ_1. So the eigenvectors

$$\begin{pmatrix} \mp 1 & -i \\ i & \mp 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\mp a - ib = 0$$

$$ib = \mp a$$

$$b = \pm ia$$

And thus

$$\hat{e}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm i \end{pmatrix}$$

Finally, for σ_3

$$\begin{vmatrix} 1 - \lambda_3 & 0 \\ 0 & -1 - \lambda_3 \end{vmatrix} = 0$$

$$(\lambda_3 - 1)(\lambda_3 + 1) = 0$$

$$\lambda_3 = \pm 1$$

Again, same as in the previous two cases. For the eigenvectors

$$\begin{pmatrix} 1 \mp 1 & 0 \\ 0 & -1 \mp 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

This gives as two equations

$$(1 \mp 1)a = 0$$

and

$$(-1 \mp 1)b = 0$$

$$(1 \pm 1)b = 0$$

These are separately satisfied if $a = (1 \pm 1)c$ and $b = (1 \mp 1)d$, where c and d are some numbers. The normalized eigenvectors are therefore

$$\hat{e}_3 = \frac{1}{2} \begin{pmatrix} 1 \pm 1 \\ 1 \mp 1 \end{pmatrix}$$

In summary, the eigenvalues of the Pauli matrices are always $\lambda = \pm 1$ and the eigenvectors are as follows

$$\hat{e}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}, \hat{e}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm i \end{pmatrix}, \hat{e}_3 = \frac{1}{2} \begin{pmatrix} 1 \pm 1 \\ 1 \mp 1 \end{pmatrix}$$

7.4   Dirac Equation in the Rest Frame of the Particle

We have determined matrices α_i and β and therefore the Dirac equation is complete as

$$\left(i \frac{\partial}{\partial t} + i\vec{\alpha} \cdot \nabla - \beta m \right) \Psi = 0$$

Since each component of $\vec{\alpha}$ and matrix β are all 4x4 matrices, the wavefunction becomes a vector with four components

$$\Psi = \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{pmatrix}$$

which is usually called the Dirac spinor. It is not shown here, but the Dirac equation can be shown to satisfy the continuity equation with flux density

$$\vec{j} = (\Psi^T)^* \vec{\alpha} \Psi$$

and probability density

$$\rho = (\Psi^T)^* \beta \Psi$$

Now, we want to find whether there are spins present in the theory. This should be visible from the model of a particle in its rest frame (and in fact, all physical properties could be determined in the rest frame and then just transformed by Lorentz transformations to any other frame). In the rest frame of the particle, the particle has no momentum, and therefore the Dirac equation becomes

$$\left(i\frac{\partial}{\partial t} - \beta m\right)\Psi = 0 \tag{55}$$

Now we try to look for plane wave solutions in a form of

$$\Psi = \vec{a}e^{-iEt} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} e^{-iEt}$$

The Dirac equation for these solutions is

$$i\frac{\partial}{\partial t}\vec{a}e^{-iEt} = m\beta\vec{a}e^{-iEt}$$

$$m\beta\vec{a}e^{-iEt} = E\vec{a}e^{-iEt}$$

$$(\beta - \frac{E}{m}I_4)\vec{a} = 0$$

Hence amplitude spinors $\vec{a}$ are the eigenvectors of β matrix. To find the eigenvalues and eigenvectors

$$|\beta - \frac{E}{m}I_4| = 0$$

$$\begin{vmatrix} -\frac{E}{m}I & I \\ I & -\frac{E}{m}I \end{vmatrix} = 0$$

$$|-\frac{E}{m}I||-\frac{E}{m}I| - |I||I| = 0$$

Since $|I| = 1$

$$\frac{E^2}{m^2} - 1 = 0$$

$$\frac{E}{m} = \pm 1$$

$$E = \pm m$$

Hence the total energy is the rest mass, but potentially negative. The negative total energy was then interpreted by Dirac as the antiparticle, which will be discussed later.

The eigenvectors satisfy

$$\begin{pmatrix} \mp I & I \\ I & \mp I \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Which leads to two separated equations

$$\mp a_1 + a_3 = 0$$

$$\mp a_2 + a_4 = 0$$

Since we do not search for trivial eigenvectors $a_1 = a_2 = a_3 = a_4 = 0$, we can construct two pairs of normalized eigenvectors

$$\vec{a}_1 = \begin{pmatrix} 1 \\ 0 \\ \pm 1 \\ 0 \end{pmatrix}$$

and

$$\vec{a}_{-1} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \pm 1 \end{pmatrix}$$

Since both $\vec{a}_1$ states have the energy equal to m, we can interpret them as the states of the normal particle in rest frame, where the two states are possible to account for the spin degeneracy - two possible spin orientations.

Similarly, $\vec{a}_{-1}$ are two states of the anti-particle with opposite spins.

And what about the negative energies? Dirac suggests that a field of possible states fills the space with state energies going approximately as $E = \sqrt{p^2 + m^2}$ for particles and $E^* = -\sqrt{p^2 + m^2}$ for anti-particle states. The reason why the particles do not simply occupy the states E^* is because we assume that all states E^* are occupied. Then, if we excite a particle from this state, a hole appears with the negative energy, that can anihilate with the original particle and has an effective positive charge. This behaviour was a good description of positron particle - the anti-particle of the electron.

However, this description requires the whole vacuum to be filled with some massive particles, which is a bit problematic. The standard model gets around this by assuming that all particles are in fact massless and their mass originates from the interaction with the Higgs field. However, massless particles do not have a rest frame, therefore they need to be treated separately.

7.5 Dirac Equation for Massless Particles

For massless particles, the Dirac equation becomes

$$(i\frac{\partial}{\partial t} + i\vec{\alpha} \cdot \nabla)\Psi = 0$$

Assuming that the solutions are plane waves of form (and using $\vec{k} \cdot \vec{r} = \frac{(\vec{p'})}{c\hbar} \cdot \vec{r} = (\vec{p'}) \cdot (\vec{r'})$)

$$\Psi = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} e^{i(\vec{p} \cdot \vec{r} - Et)}$$

we have

$$E\Psi = -i\vec{\alpha} \cdot \nabla \Psi$$

$$E\Psi = \begin{pmatrix} i\vec{\sigma} \cdot \nabla & 0 \\ 0 & -i\vec{\sigma} \cdot \nabla \end{pmatrix} \Psi$$

where

$$\vec{\sigma} = \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix}$$

This equation clearly separates into to equations, one for

$$\Psi_L = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} e^{i(\vec{p} \cdot \vec{r} - Et)}$$

and second for

$$\Psi_R = \begin{pmatrix} a_3 \\ a_4 \end{pmatrix} e^{i(\vec{p} \cdot \vec{r} - Et)}$$

as

$$E\Psi_L = i\vec{\sigma} \cdot \nabla \Psi_L$$

and

$$E\Psi_R = -i\vec{\sigma} \cdot \nabla \Psi_R$$

The first equation becomes explicitly

$$E\Psi_L = i\vec{\sigma} \cdot \nabla \Psi_L = i\left(\sigma_1 \frac{\partial}{\partial x} + \sigma_2 \frac{\partial}{\partial y} + \sigma_3 \frac{\partial}{\partial z}\right) \vec{a}_L e^{i(\vec{p} \cdot \vec{r} - Et)} = i\left(\sigma_1 ip_x + \sigma_2 ip_y + \sigma_3 ip_z\right) \vec{a}_L e^{i(\vec{p} \cdot \vec{r} - Et)}$$

$$E\vec{a}_L = -\left(\sigma_1 p_x + \sigma_2 p_y + \sigma_3 p_z\right)\vec{a}_L$$

$$E\vec{a}_L = -\left[\begin{pmatrix} 0 & p_x \\ p_x & 0 \end{pmatrix} + \begin{pmatrix} 0 & -ip_y \\ ip_y & 0 \end{pmatrix} + \begin{pmatrix} p_z & 0 \\ 0 & -p_z \end{pmatrix}\right]\vec{a}_L = -\begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix}\vec{a}_L$$

Hence

$$\left[\begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix} + EI\right]\vec{a}_L$$

which is an eigenvalue equation for eigenvalue $(-E)$. Hence

$$\begin{vmatrix} E + p_z & p_x - ip_y \\ p_x + ip_y & E - p_z \end{vmatrix} = 0$$

$$(E + p_z)(E - p_z) - (p_x - ip_y)(p_x + ip_y) = 0$$

$$E^2 = p_x^2 + p_y^2 + p_z^2$$

Hence

$$E = \pm\sqrt{p_x^2 + p_y^2 + p_z^2} = \pm|\vec{p}|$$

this has common interpretation of particle for $E = |\vec{p}|$ and anti-particle for $E = -|\vec{p}|$. The eigenvectors satisfy

$$\begin{pmatrix} p_z \pm |\vec{p}| & p_x - ip_y \\ p_x + ip_y & -p_z \pm |\vec{p}| \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$(p_x + ip_y)a_1 - (p_z \mp |\vec{p}|)a_2 = 0$$

$$a_2 = \frac{p_x + ip_y}{p_z \mp |\vec{p}|}$$

Hence

$$\vec{a}_L = \begin{pmatrix} 1 \\ \frac{p_x + ip_y}{p_z \mp |\vec{p}|} \end{pmatrix}$$

This is for $|\vec{p}| \neq p_z$. In case when $|\vec{p}| = p_z$, which implies that $p_x = p_y = 0$, we have

$$\begin{pmatrix} p_z \pm p_z & 0 \\ 0 & -p_z \pm p_z \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

And eigenvectors are $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Similarly for the second part of spinor

$$E\Psi_R = -i\vec{\sigma} \cdot \nabla\Psi_R$$

$$E\vec{a}_R = \begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix}\vec{a}_R$$

Which means that the eigenvectors are the same, but eigenvalues are switched.

The particles with $\vec{a}_R$ eigenvectors are called the right-handed particles, those with $\vec{a}_L$ are called the left handed. The particles with the negative total energy remain the anti-particles.

7.6 Forces in Standard Model

Forces in standard model are modelled as particle trasfer between interacting particles. The particles mediating this interaction are called the gauge bosons. For electromagnetic force, the gauge boson is photon. For weak interaction, there are three gauge bosons - so called vector bosons. For strong interaction, there are 8 gluons, each identical but bearing a different set of colour+anti-colour , with one combination being prohibited. We do not speak about gravity.

For us, we can use Feynmann diagrams to qualitatively describe the interactions of the particles with their gauge bosons. The quantity that we will be deriving from these diagrams is mainly the probability of certain process. This probability is always proportional to the strength of coupling of the particle to the force field (by some coupling constant) and the propagation term of the gauge boson, squared. See figure below for an example.

To estimate α, we consider the maximum energy transfered by the photon.
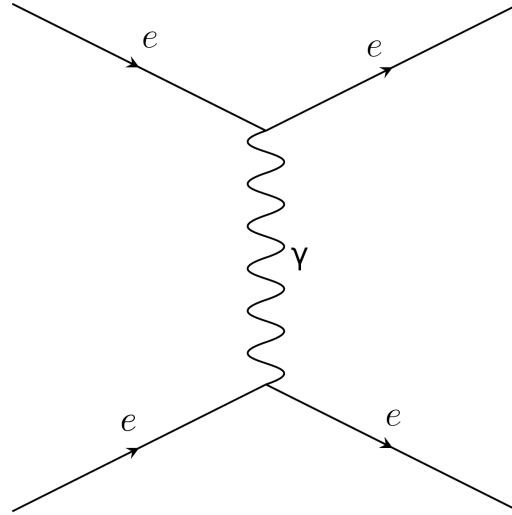
Figure 2: The Coulomb interaction - repulsion of two electrons. The probability of this process is proportional to the square of the probability amplitude of this process. The probability amplitude is proportional to the product of probability amplitudes of the separate processes in this diagram. The processes are emission of photon (proportional to coupling constant α), propagation of photon (only decreasing as a free radial wave - as $\frac{1}{r}$), and absorption of photon (again, proportional to α). Hence the probability amplitude is proportional to $\frac{\alpha^2}{r}$ and interaction probability to $\frac{\alpha^4}{r^2}$.

Since the photon is emmited from essentially nothing in the rest frame of the emmiting electron, it has to be constrained by the Heisenberg uncertainty principle (generalized) as

$$E_\gamma \Delta t \approx \hbar$$

where Δt is the time of the existence of photon, and hence the time of propagation of photon. As the photon moves at the speed of light

$$E_\gamma \frac{r}{c} = \hbar$$

$$E_\gamma = \frac{c\hbar}{r} = \frac{1}{r'}$$

Hence the coupling constant should be the ratio between the potential energy of the two electrons and the energy that would be used to exchange photon between them

$$\alpha = \frac{E_p}{E_\gamma} = \frac{e^2}{4\pi\epsilon_0 r}\frac{r}{c\hbar} = \frac{e^2}{4\pi\epsilon_0 c\hbar} \approx \frac{1}{137}$$

α is also sometimes called the fine structure constant. We have seen that we could define the energy of the photon in terms of natural units as $\frac{1}{r'}$. We could also similarly define the charge as a dimensionless number using the potential energy of the electrons as

$$\frac{(e')^2}{4\pi r'} = \frac{e^2}{4\pi\epsilon_0 r}$$

$$e' = e\sqrt{\frac{1}{\epsilon_0 \hbar c}}$$

Which also leads to

$$\alpha = \frac{(e')^2}{4\pi}$$

and interaction probability amplitude as $\frac{\alpha^2}{r'}$. From now on, I will again drop the primes.

7.6.1   Weak Force

Weak force is specific by having massive gauge bosons. This means that the energy trasfered by the gauge boson is usually smaller than its mass. This means that the boson is effectively imaginary. The radial wave of the massive boson is proportional to

$$\Psi_W \propto \frac{e^{ipr}}{r} = \frac{e^{i\sqrt{E^2-m^2}r}}{r} = \frac{e^{-\sqrt{m^2-E^2}r}}{r}$$

And for E much smaller than m

$$\Psi_W \approx \frac{e^{-mr}}{r}$$

Which is much smaller than for the electromagnetic force. In fact, the electromagnetic force can be viewed as the limit of the weak force at gauge boson mass $m \to 0$.

The coupling constant of the electroweak force is usually labeled g_2 and satisfies

$$g_2^2 = \frac{e^2}{\sin^2 \theta_W} = \frac{4\pi\alpha}{\sin^2 \theta_W}$$

This means that the weakness of the force is not in the coupling constant - the coupling constant is bigger than α. The weakness is simply in the range of the force and in the propagation term.

Experimentally, the angle θ_W could be determined from the events involving the Z^0 vector boson, which mediated some weak processes of neutral particles, and thus did not get obstructed by the electromagnetic force. This meant that the g_2 could be determined and from probabilites of β decays (which are weak force processes), the mass of the vector bosons could be established. It was later measured at CERN $p\bar{p}$ collider as 80 $\frac{\text{GeV}}{c^2}$ for $W^\pm$ and 90 $\frac{\text{GeV}}{c^2}$ for Z^0.

With this value, we can try to estimate the range of the weak force, which is dominated by the exponential term.

The range of the force in natural units is approximately

$$R' = \frac{1}{m'}$$

Hence in laboratory units

$$\frac{R}{\hbar c} = \frac{1}{mc^2}$$

$$R = \frac{\hbar}{cm} \approx 2 \times 10^{-18} \text{ m} = 2 \times 10^{-3} \text{ fm} \approx 2.4 \times 10^{-3} R_p$$

where R_p is the wavelength of the proton in the nucleus (effective radius of the proton in the nucleus). Therefore, the weak force has very small range and can be for most purposes modeled as a delta potential force.

7.6.2   Fermi Weak Delta Force

Fermi's model for the weak force states that the probability of the weak interaction happening at position $\vec{r}$ is

$$\Psi_{W,F} = 4\sqrt{2}G_F\delta^3(\vec{r})$$

where G_F is the Fermi coupling constant. We can relate this constant to the real probability amplitude of the process, integrated over all space

$$\int_V \Psi_{W,F}dV = \int_V \frac{g_2^2}{4\pi}\frac{e^{-mr}}{r}dV$$

where the 4π factor is due to normalization. Using spherical polar coordinates

$$4\sqrt{2}G_F = \int_0^\infty \frac{g_2^2}{4\pi}\frac{e^{-mr}}{r}r^2dr\int_0^\pi\int_0^{2\pi}\sin\theta d\phi d\theta = g_2^2\int_0^\infty \frac{1}{4\pi}re^{-mr}dr4\pi = g_2^2\int_0^\infty re^{-mr}dr$$

Using substitution $r = \frac{x}{m}$

$$4\sqrt{2}G_F = \frac{g_2^2}{m^2}\int_0^\infty xe^{-x}dx = \frac{g_2^2}{m^2}\left(\left[-xe^{-x}\right]_0^\infty + \int_0^\infty e^{-x}dx\right) = \frac{g_2^2}{m^2}\left[-e^{-x}\right]_0^\infty = \frac{g_2^2}{m^2}$$

Hence

$$m^2 = \frac{g_2^2}{4\sqrt{2}G_F} = \frac{4\pi\alpha}{4\sqrt{2}G_F \sin^2\theta_W} = \frac{\pi\alpha}{\sqrt{2}G_F \sin^2\theta_W}$$

This is what was also originally used to estimate the masses of vector bosons, as G_F was measured from fits of experiments and θ_W was known.

7.6.3 Strong Force

Strong force is mediated by gluons and has many perculiar properties. For example, it gets generally stronger over longer distance - it works like spring. But, since gluons them selves have a colour, the overall effect is decreasing with distance.

This force creates problems when modelling it, because we cannot use perturbation theory - cannot assume that the potential energy at infinity is zero. The model coupling constant of the strong force is

$$\alpha_s = \frac{g_s}{4\pi} = \frac{4\pi}{\left(11 - \frac{2}{3}n_f\right)\ln\left(\frac{Q^2}{\Lambda^2}\right)}$$

where $n_f = 6$ is the number of flavours of quarks, Q is the momentum transferred by the gluon (and thus distance is $\propto \frac{1}{Q}$, by HUP) and Λ is some parameter. We need coupling constant to be positive, otherwise there might be negative probabilites of events, which are not defined.

Thus we need $Q \geq \Lambda$ to be able to model it. The obscure region of the strong force than starts at distances about $\Lambda \approx 1$ fm, which is about the size of the proton.

7.7 Flavour Oscillations

A surprisingly common phenomena in particle physics is the phenomena of flavour oscillations. This is caused by particles interacting with the weak force somehow and changing the flavour as they travel through space. This behaviour can be modelled in our matrix formalism. The observable states are the flavour states of the particle, while the energy eigenstates are unknown, but the mixing hamiltonian matrix is provided as a model with some free parameters that are then fitted from the experiment.

7.7.1 K Meson Oscillations

For vector of flavour states amplitudes $\begin{pmatrix} a_{K_0} \\ a_{\bar{K}_0} \end{pmatrix}$, the hamiltonian matrix has form

$$\begin{pmatrix} m_0 - i\frac{\Gamma}{2} & m_{12} - i\frac{\Gamma_{12}}{2} \\ m_{12}^* - i\frac{\Gamma_{12}^*}{2} & m_0 - i\frac{\Gamma}{2} \end{pmatrix}$$

Which can be reexpressed as

$$\mathbf{H} = \begin{pmatrix} m_0 - i\frac{\Gamma}{2} & p^2 \\ q^2 & m - i\frac{\Gamma}{2} \end{pmatrix}$$

The eigenvalues follow from

$$\mathbf{H} = \begin{pmatrix} 0 & p^2 \\ q^2 & 0 \end{pmatrix} + \left(m_0 - \frac{\Gamma}{2}\right)I = \mathbf{A} + \left(m_0 - \frac{\Gamma}{2}\right)I$$

where I is the identity matrix. The eigenvalues of $\mathbf{A}$ satisfy

$$\begin{vmatrix} -\lambda & p^2 \\ q^2 & -\lambda \end{vmatrix} = \lambda^2 - p^2q^2 = 0$$

$$\lambda = \pm pq$$

Hence the eigenvalues of $\mathbf{H}$ are $m_0 - \frac{\Gamma}{2} \pm pq$. The eigenvectors satisfy

$$\begin{pmatrix} \mp pq & p^2 \\ q^2 & \mp pq \end{pmatrix} \begin{pmatrix} b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\mp pqb + p^2c = 0$$

$$c = \pm\frac{q}{p}b$$

Hence, the normalized eigenvectors are

$$\hat{b}_{\pm} = \frac{1}{\sqrt{|p|^2 + |q|^2}} \begin{pmatrix} p \\ \pm q \end{pmatrix}$$

The Schrödinger equation is

$$i\frac{\partial}{\partial t} \begin{pmatrix} a_{K_0} \\ a_{\bar{K}_0} \end{pmatrix} = \begin{pmatrix} m_0 - \frac{\Gamma}{2} & p^2 \\ q^2 & m_0 - \frac{\Gamma}{2} \end{pmatrix} \begin{pmatrix} a_{K_0} \\ a_{\bar{K}_0} \end{pmatrix}$$

These vectors can be rewritten as

$$\begin{pmatrix} a_{K_0} \\ a_{\bar{K}_0} \end{pmatrix} = a_{K_0} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + a_{\bar{K}_0} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Here

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{\sqrt{|p|^2 + |q|^2}}{2p} \left[\frac{1}{\sqrt{|p^2| + |q|^2}} \left(\begin{pmatrix} p \\ q \end{pmatrix} + \begin{pmatrix} p \\ -q \end{pmatrix} \right) \right] = \frac{\sqrt{|p|^2 + |q|^2}}{2p} \left(\hat{b}_+ + \hat{b}_- \right)$$

Similarly

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{\sqrt{|p|^2 + |q|^2}}{2q} \left(\hat{b}_+ - \hat{b}_- \right)$$

And thus

$$\begin{pmatrix} a_{K_0} \\ a_{\bar{K}_0} \end{pmatrix} = \sqrt{|p|^2 + |q|^2} \left(\frac{a_{K_0}}{2p}(\hat{b}_+ + \hat{b}_-) + \frac{a_{\bar{K}_0}}{2q}(\hat{b}_+ - \hat{b}_-) \right)$$

$$\begin{pmatrix} a_{K_0} \\ a_{\bar{K}_0} \end{pmatrix} = \sqrt{|p|^2 + |q|^2} \left(\frac{a_{K_0}}{2p} + \frac{a_{\bar{K}_0}}{2q} \right) \hat{b}_+ + \sqrt{|p|^2 + |q|^2} \left(\frac{a_{K_0}}{2p} - \frac{a_{\bar{K}_0}}{2q} \right) \hat{b}_-$$

Hence the Schrödinger equation is

$$i\frac{\partial}{\partial t} \left(\frac{a_{K_0}}{2p} + \frac{a_{\bar{K}_0}}{2q} \right) \hat{b}_+ + i\frac{\partial}{\partial t} \left(\frac{a_{K_0}}{2p} - \frac{a_{\bar{K}_0}}{2q} \right) \hat{b}_- =$$

$$= \left(\frac{a_{K_0}}{2p} + \frac{a_{\bar{K}_0}}{2q} \right) \left(m_0 - i\frac{\Gamma}{2} + pq \right) \hat{b}_+ + \left(\frac{a_{K_0}}{2p} - \frac{a_{\bar{K}_0}}{2q} \right) \left(m_0 - i\frac{\Gamma}{2} - pq \right) \hat{b}_-$$

Since the eigenvectors of energy are constant in time, we can separate this equation into two equations for the amplitudes. The first (for $\hat{b}_+$)

$$i\frac{\partial}{\partial t} \left(\frac{a_{K_0}(t)}{2p} + \frac{a_{\bar{K}_0}(t)}{2q} \right) = \left(m_0 - i\frac{\Gamma}{2} + pq \right) \left(\frac{a_{K_0}(0)}{2p} + \frac{a_{\bar{K}_0}(0)}{2q} \right)$$

Hence

$$\left(\frac{a_{K_0}(t)}{2p} + \frac{a_{\bar{K}_0}(t)}{2q} \right) = \left(\frac{a_{K_0}(0)}{2p} + \frac{a_{\bar{K}_0}(0)}{2q} \right) e^{\left(-im_0 - ipq - \frac{\Gamma}{2}\right)t}$$

And similarly

$$\left(\frac{a_{K_0}(t)}{2p} - \frac{a_{\bar{K}_0}(t)}{2q} \right) = \left(\frac{a_{K_0}(0)}{2p} - \frac{a_{\bar{K}_0}(0)}{2q} \right) e^{\left(-im_0 + ipq - \frac{\Gamma}{2}\right)t}$$

Now imagine that the particle was prepared as K_0, i. e. $a_{\bar{K}_0} = 0$. Then

$$\frac{a_{K_0}}{p} = \frac{a_{K_0}}{p} e^{\left(-im_0 - \frac{\Gamma}{2}\right)t} \left(e^{-ipqt} + e^{ipqt} \right)$$

And hence the probability of observing the particle at sime time t is

$$a_{K_0}(t)a_{K_0}^*(t) = |a_{K_0}(0)|^2 e^{-\Gamma t} \left(e^{ipqt} + e^{-ipqt} \right) \left(e^{-ip^*q^*t} + e^{ip^*q^*t} \right)$$