Junior Honours Quantum Mechanics

Chris Stock (cstock@ed.ac.uk)

September 18, 2017

1 Introduction

1.1 Outline

This course will involve a short introduction on the breakdown of classical theory followed by a series of lectures on the wave mechanics approach to quantum theory. The second part of the course will discuss Dirac notation and applications of this using matrix mechanics. Understanding the techniques applied to matrix mechanics is key to this course.

The outline of the class is listed below.

1) Introduction- what is wrong with classical theory?
2) Wave mechanics - introduction, "postulates", Schrodinger equation, and introduction to angular momentum
3) Applications to scattering
4) The Hydrogen atom
5) Simple Harmonic Oscillator solved using Hermite polynomials
6) Matrix mechanics - Dirac notation
7) Observables and operator algebra (CSCO-Complete Set of Commuting Observables)
8) Ladder operators and the harmonic oscillator revisited
9) Angular momentum
10) Spin operators
11) The two level problem

1.2 References

[1] C. Cohen-Tannoudji, D. Diu, and F. Laloe, *"Quantum Mechanics"* (Hermann, Paris, France, 1973).
(very formal treatment of the mathematical basis for quantum mechanics with emphasis on matrix mechanics)

[2] L. Pauling, and E.B. Wilson, *"Introduction to Quantum Mechanics with Applications to Chemistry"* (Dover, New York, 1935).
(good presentation of the solution to the hydrogen atom, perturbation methods, and the harmonic oscillator using wave mechanics)

[3] M.A. Morrison, T.L. Estle, and N.F. Lane, *"Quantum States of Atoms, Molecules, and Solids"* (Prentice-Hill, New Jersey, 1976).
(excellent presentation and discussion of some key aspects to atomic physics and also perturbation methods)

[4] B. R. Judd, *"Operator Techniques in Atomic Spectroscopy"* (Princeton University Press, New Jersey, 1998).
(a short book providing a compact discussion of angular momentum and operators for crystal field

analysis)

[5] L. D. Landau and E. M. Lifshitz, *"Quantum Mechanics"* (Elsevier Science, Oxford, 1977).
(advanced treatment of quantum mechanics with little use of Dirac Notation)

[6] A. Messiah, *"Quantum Mechanics"* (Dover, New York, 1958).
(a very good treatment of both matrix and wave mechanics)

[7] J. Hardy, *"Quantum Physics"* (Notes from second year quantum physics).
(These notes provide a starting point for this course and some reference will be made. The two text-books used to formulate these notes by French and Taylor and Eisberg also provide useful starting references.)

[8] J. A. Peacock *"Fourier Analysis"* (Notes from current third year course).
(The techniques developed in this course will be referred to and are fundamental to many of the results described here.)

2 What is wrong with classical mechanics? (Messiah -Chapter 1, Cohen-Tannoudji-Chapter 1)

There are many examples and illustrations where experiments clearly show that classical theory is incomplete and something else is required to describe real systems. Probably most people are familiar with the "UV catastrophe" problem and the photoelectric effect (summarized by the relation $E = h\nu = \hbar\omega$) which clearly show the quantized nature of light (see notes on the photoelectric effect by Hardy). In parallel to these, the discovery of fundamental particles such as the electron highlight the quantized nature of matter. It therefore is clear that a new theory is needed and must treat matter and light on an equal footing.

While these experiments highlight the failure of continuous classical theory, I would like to focus on two experiments that describe what properties should exist in a "new theory".

2.1 Experiment 1: Double slit experiment

There are several conclusions we can draw from this experiment. First, while the photoelectric effect tells us that light should be considered in terms of discrete units called photons, the interference effect illustrates the underlying wave nature. The theory that we develop must have "wave-particle duality". While most people are familiar with the double slit experiment resulting in interference patterns (Fig. 1) with light, a similar effect can be observed with particles such as electrons. An electron diffraction pattern is shown in Fig. 2 which similar to that of light. This indicates that electrons can be behave qualitatively like waves, while the photoelectric effect shows that light can behave like particles.

2.1.1 de Broglie relation

We have shown above the need for both a wave and particle description of light based upon our knowledge of the photoelectric effect (see J. Hardy's notes) and Young's double slit experiment. The wave particle duality is formalized through the de Broglie relation which is written as follows.

$$\lambda = \frac{2\pi}{k} = \frac{h}{p} \tag{1}$$

The energy of the particle is related to the frequency by the Einstein relation.

$$E = \hbar\omega \tag{2}$$

Recall that $\omega \equiv 2\pi\nu$.

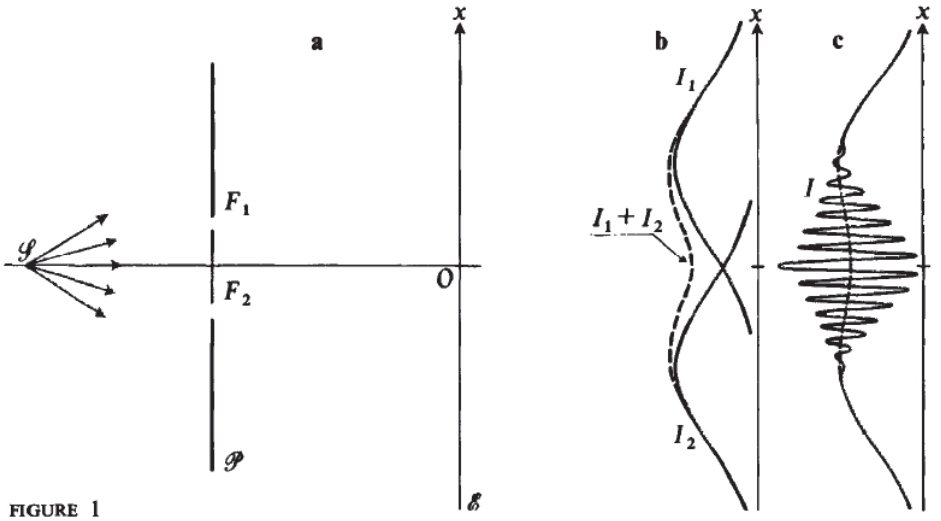

FIGURE 1

Diagram of Young's double-slit light interference experiment (fig. a). Each of the slits F_1 and F_2 produces a diffraction pattern on the screen $\mathscr{E}$. The corresponding intensities are $I_1(x)$ and $I_2(x)$ (solid lines in figure b). When the two slits F_1 and F_2 are open simultaneously, the intensity $I(x)$ observed on the screen is not the sum $I_1(x) + I_2(x)$ (dashed lines in figures b and c), but shows oscillations due to the interference between the electric fields radiated by F_1 and F_2 (solid line in figure c).

Figure 1: Figure taken Cohen-Tannoudji, *Quantum Mechanics*.

2.1.2 Experiment 2: Malus' law

Consider an experiment consisting of a polarized plane monochromatic light wave incident on a detector.

Classically, the light wave takes the following form,

$$\vec{E}(\vec{r},t) = E_\circ e^{i(kz-\omega t)}\hat{p} \tag{3}$$

After passing through the analyzer, the plane wave is polarized along a particular direction,

$$\vec{E}'(\vec{r},t) = E'_\circ e^{i(kz-\omega t)}\hat{x} \tag{4}$$

The intensity of this transmitted wave is proportional to $|E'_\circ|^2$ and is given by Malus' law - $I' = I\cos^2\theta$ where θ is the angle between $\hat{x}$ and $|\hat{p}|$. This law is difficult to understand at a quantum level when the incident beam is so weak that the photon will either cross the polarization analyzer or be absorbed (we cannot detect fractions of photons). The result does make sense if we consider a large number of photons incident on the analyzer where the detected intensity will be proportional to $N\cos^2\theta$ photons. There are several results.

The analyzer only chooses particular chosen or "eigen" (we will discuss more why we use this word later) results. This makes sense if we think about the analyzer only pointing along two possible directions, but if it is an angle (like in Fig. 3) we can think about the result in terms of probabilities. This experiment highlights the need to consider probabilities in the new theory and to find these probabilities, we decomposes the particles into a linear combination of various eigenstates or chosen states. This idea is often referred to as spectral decomposition.

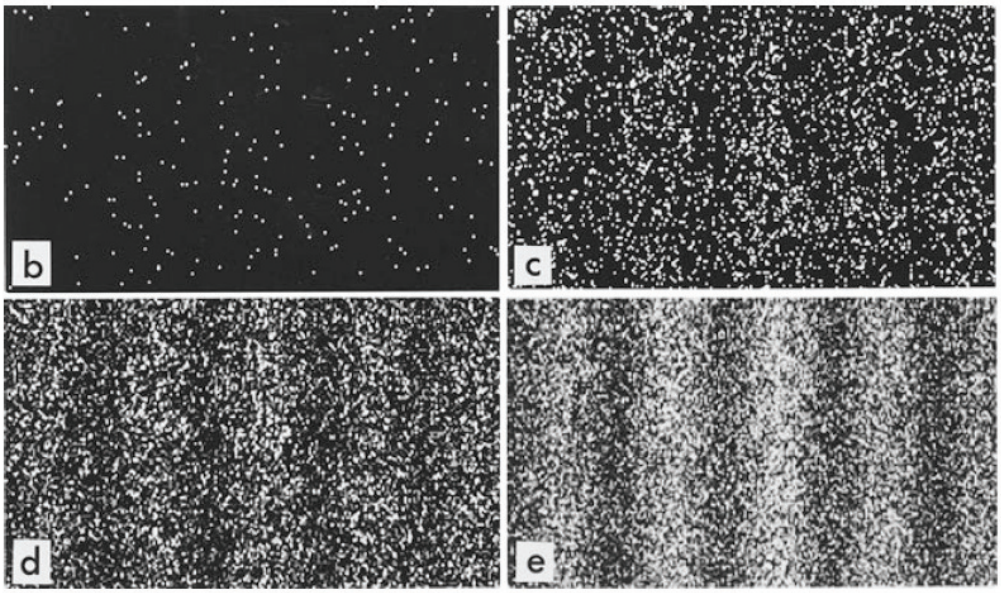

FIGURE 1.4. Four images from the 1989 experiment at Hitachi showing the impact of individual electrons gradually building up to form an interference pattern. Image by Akira Tonomura and Wikimedia Commons user Belsazar. File is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license.

Figure 2: Figure taken from B. C. Hall, *Quantum Theory for Mathematicians*.

The final interesting result from this experiment is that the measurement disturbs the system being probed. We will come back to this point later. A final point that needs to be made is that measuring device is linear in terms of E and note that $I \propto E^2$. This is particularly important in the context of the double slit experiment.

2.2 Correspondence principle and cleaning up classical mechanics

In the discussion above we have discussed two examples (Young's double slit experiment and Malus' law) where classical theory seems to be incompatible with experimental results. However, classical theory does work (and is very successful) so our new quantum theory must be equivalent to classical theory in some limit. A natural question to ask is - Can we clean up classical mechanics by noting that in the "classical limit" our new theory should reduce the familiar classical results?

3 Wave mechanics - postulates and the Schrodinger equation

The main conclusion from the experiments discussed in the previous section that a new theory must be linear and obey the superposition principle. Note the similarity between these physical constraints and the definitions above for a linear operator acting on a vector.

Lets describe some postulates that we have learned from these two experiments to create a formulation of quantum mechanics:

(i) "The quantum state of a particle such as the electron is characterized by a wave function $\psi(\vec{r}, t)$, which contain all of the information it is possible to obtain about the particle."

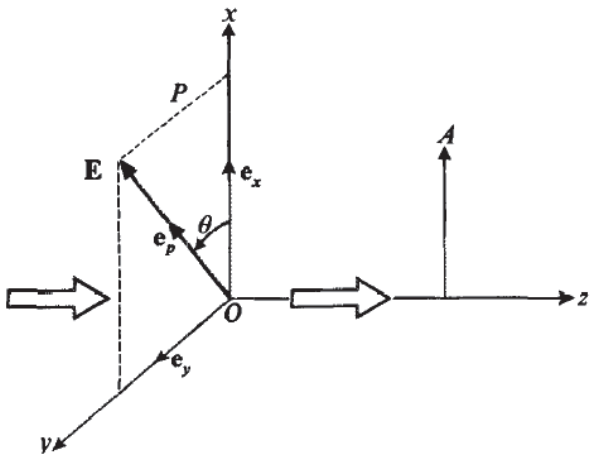

FIGURE 2

A simple measurement experiment relating to the polarization of a light wave. A beam of light propagates along the direction Oz and crosses successively the polarizer P and the analyzer A. θ is the angle between Ox and the electric field of the wave transmitted by P. The vibrations transmitted by A are parallel to Ox.

Figure 3: Figure taken from Cohen-Tannoudji, *Quantum Mechanics, vol 1.*

(*ii*) "$\psi(\vec{r},t)$ is interpreted as a probability amplitude of the particle's presence. Since the possible positions of the particle form a continuum, the probability $dP(\vec{r},t)$ of the particle being at time t, in a volume element $d^3r = dxdydz$ situated at the point $\vec{r}$ must be proportional to d^3r and therefore infinitesimal. $|\psi(\vec{r},t)|^2$ is then interpreted as the corresponding probability density with $dP = C|\psi(\vec{r},t)|^2 d^3r$, where C is a normalization constant."

(*iii*) "*The principle of decomposition* applies to the measurement of an arbitrary physical quantity: 1) The result found must belong to a set of eigen results a; 2) With each eigenvalue a is associated an eigenstate, that is, an eigenfunction $\psi(\vec{r})$. This function is such that, if $\psi(\vec{r},t_0) = \psi_a(\vec{r})$ (where t_0 is the time at which the measurement is performed), the measurement will always yield a; 3) For any $\psi(\vec{r},t)$, the probability P_a of finding the eigenvalue a for a measurement at the time t_0 is found by decomposing $\psi(\vec{r},t)$ in terms of the functions $\psi_a(\vec{r})$:"

$$\psi(\vec{r},t_0) = \sum_a c_a \psi_a(\vec{r}) \tag{5}$$

then,

$$P_a = \frac{|c_a|^2}{\sum_a |c_a|^2} \tag{6}$$

(the presence of the denominator ensures that the total probability is equal to 1)

(*iv*) The equation describing the evolution of the function $\psi(\vec{r},t)$ remains to be written. It is possible to introduce it in a very nature way, using the Planck and de Broglie relations. Based on the experiments discussed above, the equation must be linear and homogeneous and hence we write the following,

$$\hat{H}\psi = E\psi \tag{7}$$

Given the correspondence principle discussed above, we can postulate what the form of this equation is based on the result that any theory we write must cross over to the classical limit and also must be obeyed by plane wave solutions of the form $\psi(\vec{r},t) \propto e^{-i\frac{E}{\hbar}t}e^{ikx}$. We therefore expect the time dependence to be linear and equal to $E \to i\hbar\frac{\partial}{\partial t}$. Also, given that classically we expect the energy of a free particle of momentum p to be $\frac{p^2}{2m}$ which (using the de Brogile relations) equal to $\frac{\hbar^2k^2}{2m}$ we therefore make the association $p \to -i\hbar\nabla$. Therefore, the Schrodinger equation for a free particle would be

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\vec{r},t) = i\hbar\frac{\partial}{\partial t}\psi(\vec{r},t) \tag{8}$$

If we have a particle in a potential $V(\vec{r},t)$, then classically we would write the energy as the sum of the kinetic energy and the potential energy. Adding this term in gives us the full Schrodinger equation.

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\vec{r},t) + V(\vec{r},t)\psi(\vec{r},t) = i\hbar\frac{\partial}{\partial t}\psi(\vec{r},t) \tag{9}$$

This is the full equation which we will be solving for a series of cases during the wave mechanics part of this course. To motivate this equation, we have associated various classical observable quantities with operations based on the correspondence principle requiring us to retain the solutions found with classical dynamics. For example, we associate momentum p of a particle with the operation $-i\hbar\nabla$ and the energy with the time derivative $i\hbar\frac{\partial}{\partial t}$.

4 Particles in a time independent scalar potential

In this section, we will solve the Schrodinger equation for several one-dimensional potential wells and establish a procedure for solving these problems in general. In the previous section, we argued that our new theory must be based upon a linear homogeneous equation based upon superposition arguments (Young's experiment) and spectral decomposition (Malus' Law). We then invoked the fact that any theory we write should give the classical result as a particular limit and also we noted that it should be consistent with our knowledge of plane wave solutions based on the photoelectric effect and classical electromagnetic theory. We then wrote the following wave equation (Schrodinger's equation).

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r},t) = -\frac{\hbar^2}{2m}\nabla^2\psi(\vec{r},t) + V(\vec{r})\psi(\vec{r},t) \tag{10}$$

In this section, we will study solutions of this wave equation for certain "toy" models including the infinitely deep well, the Dirac well, and the harmonic well. We will then go onto apply these techniques to solving the hydrogen atom.

For the remainder of this section on wave mechanics, we will be studying potentials that are *time independent* implying that V is only a function of $\vec{r}$ as written above. Therefore, using the separation of variables method for solving different equations we can start of with the following solution of Eqn. 12.

$$\psi(\vec{r},t) = \phi(\vec{r})e^{-i\frac{E}{\hbar}t} \tag{11}$$

substituting this back into the Schrodinger equation above gives the time independent Schrodinger equation

$$E\phi(\vec{r}) = -\frac{\hbar^2}{2m}\nabla^2\phi(\vec{r}) + V(\vec{r})\phi(\vec{r}) \tag{12}$$

Solving this for a series of potentials will be the topic of the next few sections.

4.1 Square well with infinite walls (Messiah-Chapter 3, Hardy - Chapter 9.1.2)

To start our study of special potentials applying Quantum wave mechanics, we start with the infinite well. This was done previously in J. Hardy's course, but it is instructive to review the potential and the techniques for solving it before studying other problems.

$$V(x) = \begin{cases} 0 & \text{if } |x| < \frac{a}{2} \\ \infty & \text{if } |x| \geq \frac{a}{2} \end{cases}$$

We can start off by noting that solution of ψ for $|x| \geq a/2$ is 0. There is no possibility that that the particle can be found within the walls. Within the region $|x| < a/2$, the Schrodinger equation can be written as,

$$\frac{d^2\psi(x)}{dx^2} + \epsilon\psi = 0 \tag{13}$$

The solution has been worked out by Hardy and we will not go into the details here but rather to state that it is,

$$\psi(x)_n = \sin\left(\frac{n\pi}{a}x\right), n = even \tag{14}$$

$$\psi(x)_n = \cos\left(\frac{n\pi}{a}x\right), n = odd \tag{15}$$

$$\tag{16}$$

with energy eigenvalues of

$$\epsilon_n = \frac{n^2\pi^2}{a^2} \tag{17}$$

Note that when $n = 0$ there is no particle in the box. This is a valid solution of the Schrodinger equation above and corresponds to zero energy. When there is a particle in the box it must have a ground state energy given by $n = 1$ in the above formula. We will return to this concept of zero point energy later in the harmonic oscillator.
In the exercises, you will be investigating the finite depth well and connecting the solution with the Dirac potential discussed in the next section.

4.2 Some special functions- the Dirac Delta and Heaviside functions(Cohen-Tannoudji-Appendix-2)

For the remainder of this course, we will require some understanding of several functions with unique properties. In particular we will make use of the "δ-function. As pointed out by Cohen-Tannoudji, this is not mathematically strictly a function but rather a distribution. We will refer to it as a function as dictated by convention.

Consider a function $\delta^\epsilon(x)$ with the following definition

$$\delta^\epsilon(x) = \begin{cases} \frac{1}{\epsilon} & \text{if } \frac{-\epsilon}{2} < x < \frac{\epsilon}{2} \\ 0 & \text{if } |x| > \frac{\epsilon}{2}. \end{cases}$$

If we consider the integral $\int_{-\infty}^{\infty} \delta^\epsilon f(x)dx \sim f(0)\int_{-\infty}^{\infty} \delta^\epsilon(x)dx = f(0)$. The smaller ϵ, the better this approximation. We therefore examine $\lim_{\epsilon\to 0}$ and define the δ function by the relation $\int_{-\infty}^{+\infty} f(x)\delta(x)dx = f(0)$. In general we can write the following,

$$\int_{-\infty}^{+\infty} \delta(x-\alpha)f(x)dx = f(\alpha). \tag{18}$$

It follows from this that $\int_{-\infty}^{\infty} \delta(x)dx = 1$. There are a number of ways of defining the δ function and for simplicity we have chosen $1/\epsilon$. Other examples which approach this when ϵ approaches zero are

$\frac{1}{2\epsilon}e^{-|x|/\epsilon}$, $\frac{1}{\pi}\frac{\epsilon}{x^2+\epsilon^2}$, and $\frac{1}{\pi}\frac{\sin(x/\epsilon)}{x}$. A more comprehensive list can be found in Cohen-Tannoudji (volume 2 and appendix 2).

Another definition of the δ-function can be formulated in terms of the step or "Heaviside" function.

$$\Theta(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x < 0. \end{cases}$$

It can be shown, taking our definition of δ above in terms of ϵ, that δ-function is the derivative of the Heaviside function. These two "functions" are very important and will be used in our study of potentials. Further properties of the δ function will be derived in the problem set.

4.3 Dirac well (Davies-Chapter 3, Cohen-Tannoudji- Complement K1)

A very interesting potential which has many applications in scattering and also molecular physics is the Dirac potential. This is also a useful potential for establishing techniques in analyzing any problem dealing with a potential well.

$$V(x) = -\alpha\delta(x) \tag{19}$$

where α is a positive real number. The Schrodinger equation then becomes the following

$$\frac{-\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} - \alpha\delta(x)\psi(x) = E\psi(x) \tag{20}$$

First consider the two limiting cases where $x \neq 0$ and hence the equation reduces to $\frac{-\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} = E\psi(x)$. The solution to this is,

$$\psi(x) = Ae^{-\beta x} + Be^{+\beta x}, x \neq 0 \tag{21}$$

where A and B need to be determined and $\beta = \sqrt{-2mE}/\hbar$. Imposing well defined properties of ψ (ie finite at all values of x) fixes the solution to be the following

$$\psi(x) = \begin{cases} Be^{\beta x} & \text{if } x < 0 \\ Ae^{-\beta x} & \text{if } x > 0. \end{cases}$$

and imposing the fact that ψ must be continuous gives us $A = B$ (take the limit as $x \to 0$ from both positive and negative sides and set them equal). We now deal with the x=0. At this point, we have the above ordinary differential equation for $\psi(x)$ containing the special function $\delta(x)$. Integrating with a region of $\pm\epsilon$ around $x = 0$ gives,

$$\int_{-\epsilon}^{+\epsilon}\left(\frac{-\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} - \alpha\delta(x)\psi(x)\right)dx = \int_{-\epsilon}^{+\epsilon} E\psi(x)dx \tag{22}$$

and integrating

$$\frac{-\hbar^2}{2m}\left(\lim_{\epsilon\to 0^+}\frac{d\psi(x)}{dx} - \lim_{\epsilon\to 0^-}\frac{d\psi(x)}{dx}\right) - \alpha\psi(0) = E\int_{-\epsilon}^{+\epsilon}\psi(x)dx \tag{23}$$

Lets now remember that $\psi(x)$ must be continuous (as well as being finite as used above already). The continuity property of $\psi(x)$ implies that $\lim_{|\epsilon|\to 0}\int_{-\epsilon}^{+\epsilon}\psi(x)dx \equiv 0$ (identically). Therefore, the right hand part of the equation above is 0. We now have,

$$\left(\lim_{\epsilon\to 0^+}\frac{d\psi(x)}{dx} - \lim_{\epsilon\to 0^-}\frac{d\psi(x)}{dx}\right) = -\frac{2m\alpha}{\hbar^2}\psi(0) \tag{24}$$

Using our solutions for $\psi(x)$ at $|x| > 0$, we find that by taking derivatives that $2A\beta = \frac{2m\alpha}{\hbar^2}\psi(0)$, and continuity implies $\lim_{x\to 0}\psi(x) = A$ implying $\beta = \frac{m\alpha}{\hbar^2}$. There is there only one bound state, or energy eigenvalue, to this potential, namely

$$E = -\frac{m\alpha^2}{2\hbar^2}. \tag{25}$$

To complete this off we still need to find A and this can be derived by noting $\int_{-\infty}^{+\infty}\psi(x)dx \equiv 1$. Doing the algebra gives $A = \frac{\sqrt{m\alpha}}{\hbar}$. We can then write the solution of ψ compactly as

$$\psi(x) = \frac{\sqrt{m\alpha}}{\hbar}e^{-\frac{m\alpha}{\hbar^2}|x|} \tag{26}$$

This solution can also be derived by taking a well of finite width and then taking the width to zero studying how many energy eigenvalues there are as a function of width.

4.4 Harmonic Oscillator (Pauling - Chapter 3, Landau- Chapter 23)

Consider a potential energy of the form $V(x) = \frac{1}{2}m\omega_\circ^2 x^2$ where x is the displacement of the particle of mass m from the equilibrium position $x = 0$. The classical energy of this system is written as follows,

$$E = \frac{p^2}{2m} + \frac{1}{2}m\omega_\circ^2 x^2. \tag{27}$$

Using the correspondence principle discussed above, we can associate E with the Hamiltonian operator H and the moment p with its corresponding operator $-i\hbar\frac{d}{dx}$. We can then write the Schrodinger equation as follows,

$$-\frac{d^2\psi(x)}{dx^2} + \alpha^2x^2\psi(x) = \lambda\psi(x) \tag{28}$$

or

$$\frac{d^2\psi(x)}{dx^2} + \left(\lambda - \alpha^2x^2\right)\psi(x) = 0. \tag{29}$$

We would like to find solutions for $\psi(x)$ that are physical and are suitable everywhere (ie for x in the range of $-\infty$ to $+\infty$). We could use our knowledge of special functions to solve this, however it is useful to approach this from a different standpoint starting from finding the solution in certain limits and then perform a power series expansion. We are searching for finite solutions which are single valued over this region (well behaved functions). First, consider the situation with $|x|$ is large in the limit where α^2x^2 above dominates and λ is negligible by design. In this limit we want to find the solution to the following equation,

$$\frac{d^2\psi(x)}{dx^2} = \alpha^2x^2\psi(x). \tag{30}$$

We can guess the solution to this equation as $\psi = e^{\pm\frac{\alpha}{2}x^2}$ and this can be verified by direct substitution and also remembering we are studying the solution at large $|x|$ (try to verify this and remember that we are working in the limit of large x). We note that of the two solutions $e^{+\frac{\alpha}{2}x^2}$ is not satisfactory given the fact that it diverges. Using the fact that we found a good solution at large $|x|$, we can now proceed to derive solutions by introducing a power series in x and deriving the coefficients.

Define a trial wavefunction $\psi(x) = e^{-\frac{\alpha}{2}x^2}f(x)$. Lets list the derivatives here,

$$\frac{d\psi}{dx} = -\alpha x f \exp\left[-\frac{\alpha}{2}x^2\right] + f' \exp\left[-\frac{\alpha}{2}x^2\right] \tag{31}$$

$$\begin{aligned}\frac{d^2\psi}{dx^2} &= \left(-\alpha f + \alpha^2 x^2 f - \alpha x f' + f'' - \alpha x f'\right) \exp\left[-\frac{\alpha}{2}x^2\right] \\ &= \left(\alpha^2 x^2 f - \alpha f - 2\alpha x f' + f''\right) \exp\left[-\frac{\alpha}{2}x^2\right]\end{aligned} \tag{32}$$

Substituting this into Eqn. 30 gives the following differential equation for f

$$f'' - 2\alpha x f' + (\lambda - \alpha) f = 0 \tag{33}$$

To make the connection with special functions in a few steps lets make the substitution $\xi = \sqrt{\alpha}x$ and replace the function $f(x)$ with $H(\xi)$ to get the following differential equation,

$$\frac{d^2H}{d\xi^2} - 2\xi\frac{dH}{d\xi} + \left(\frac{\lambda}{\alpha} - 1\right) H = 0 \tag{34}$$

We now represent $H(\xi)$ as a power series

$$H(\xi) = \sum_n a_n \xi^n = a_0 + a_1\xi + a_2\xi^2 + ... \tag{35}$$

At this point we have made an important assumption regarding the properties of H as we have assumed that this series converges for all ξ and that the function is well defined. Physically, we know this has to be the case for the solutions to make sense and to satisfy our overall constraint that the integral of the probability density must be finite : $\int_{-\infty}^{\infty} |\psi(x)|^2 dx = 1$.

By now differentiating $H(\xi)$ and substituting the derivatives back into Eqn. 34 we come up with a rather messy expression which contains powers of ξ. However, for this equation to hold for all ξ, we can equate coefficients of terms with the same power of ξ and hence derive the following recursion relation (this will be done more explicitly in the problem set). We therefore get the following recursion formula,

$$(n+1)(n+2)a_{n+2} + \left(\frac{\lambda}{\alpha} - 1 - 2n\right) a_n = 0. \tag{36}$$

We are basically done as we have derived a general solution $\psi(x)$ for the harmonic oscillator in terms of functions of the form $\psi(x) = f(x)e^{-\frac{\alpha}{2}x^2}$ with $f(x)$ being a power series. There is however a serious problem. While we assumed a few steps ago that the series $f(x)$ converged for all x, this is not generally true in the recursion relation in Eqn. 36 for all values of λ (the energy values) and α (potential well curvature). The function is square integrable if $f(x)$ is truncated and does not contain an infinite number of terms in its series expansion. Note that the constraint that we have imposed on $\psi(x)$ is much more strict than just convergence, we require the following to be true for our solutions.

$$\int_{-\infty}^{\infty} |\psi(x)|^2 dx = 1. \tag{37}$$

It can be seen from Eqn. 36 that the series expansion is finite (and hence converges $\forall x$) if

$$\left(\frac{\lambda}{\alpha} - 1 - 2n\right) = 0 \tag{38}$$

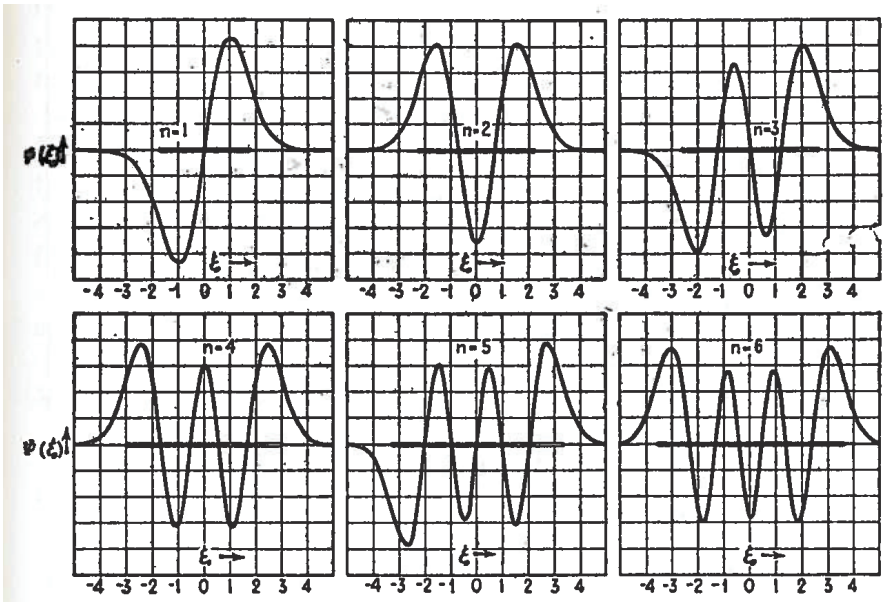

Figure 4: Figure taken Pauling, *Introduction to Quantum Mechanics.*

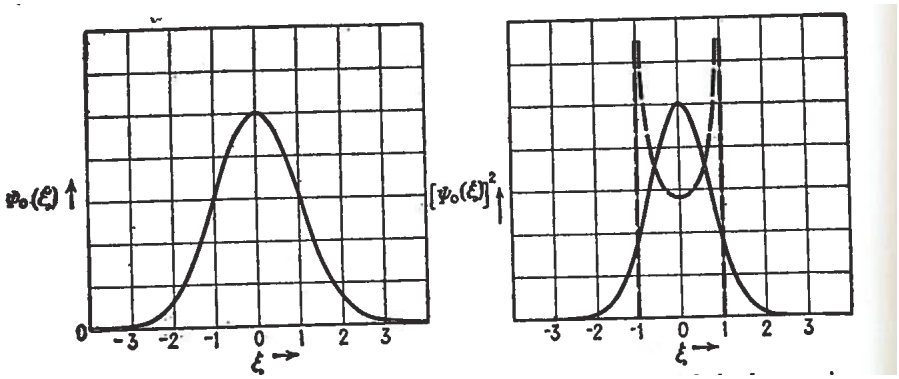

Figure 5: Figure taken Pauling, *Introduction to Quantum Mechanics.*

or the energy values have the following discrete spectra,

$$\lambda = (2n+1)\alpha. \tag{39}$$

This condition can be expressed as the following,

$$E_n = \left(n + \frac{1}{2}\right) h\nu_\circ \tag{40}$$

and this should be compared with the result derived from the photoelectric effect which states that $E = nh\nu$. The big difference here is the presence of a zero point energy when $n = 0$. This result has experimentally been confirmed and manifests as a weak attraction between two parallel plates in a vacuum. This effect only results from "vacuum" fluctuations and is known as the Casimir effect after the person who predicted it in 1948 (see Phys. Rev. Foc. **2** 28 (1998) for a nice discussion and introduction).

Graphically, examples of the wavefunctions are plotted in Fig. 4 and 6. It can be seen that the number of zeroes (or nodes) in the wavefunctions increases with energy and n with the ground state plotted in Fig. 5. An extreme case is plotted for $n = 10$ in Fig. 6.

One final note is that we have avoided using special functions here and instead of relied on the postulates of our new quantum wave theory which includes imposing well defined properties on $\psi(x)$

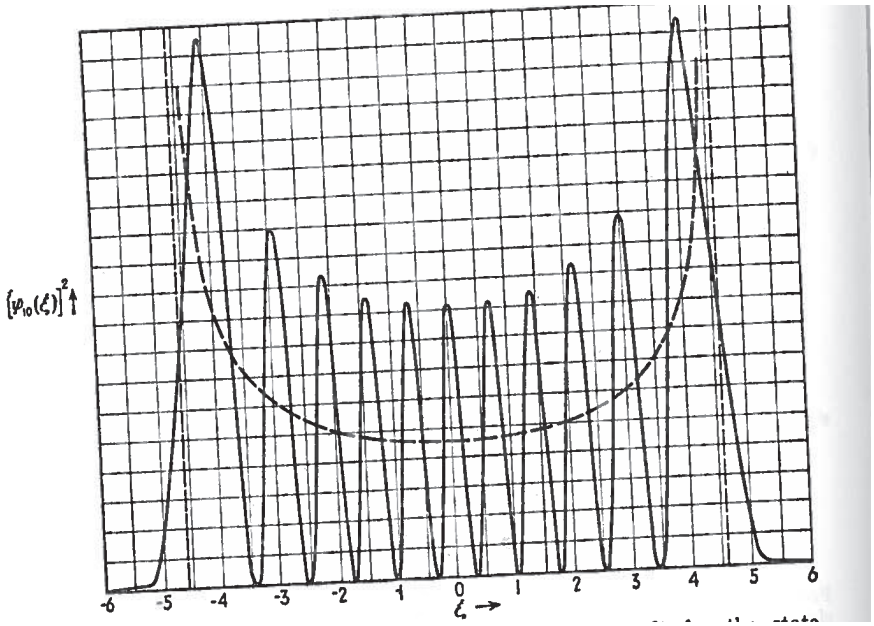

Figure 6: Figure taken Pauling, *Introduction to Quantum Mechanics.*

(like square integrable). The functions we have written down $H(\xi)$ are known as Hermite polynomials and the properties of these will be studied in problem sets or they can be investigated by reading the references noted above. In summary, using the solutions in the limit of large $|x|$ we were able to derive a general solution for the harmonic oscillator. By stipulating the physical criterion that the wavefunction must be finite and single valued, we found that naturally we had to impose a cutoff in the series expansion which resulted in the presence of discrete energy values. In the end, we found the presence of a zero point energy value and have noted that this agrees with experiments.

We will return to the harmonic oscillator problem later and solve it using operator techniques and introduce ladder operators important for studies in quantum optics and field theory.

5 Operators and angular momentum (Landau-Chapter 4)

Before solving the hydrogen atom we need to develop some formalism to describe and understand angular momentum. As noticed in classical theory, the symmetry of central force problems (where the potential only depends on the scalar r) angular momentum plays an important role because it is a conserved quantity. In classical theory, angular momentum is defined as,

$$\vec{L} = \vec{r} \times \vec{p} \tag{41}$$

It is helpful, particularly for the hydrogen atom, to first consider the classical case of a rigid rotor. Classically, a fixed particle of mass m with a fixed distance r has the following expression for the energy.

$$E = \frac{L^2}{2mr^2} \tag{42}$$

Recall from our formulation of the Schrodinger equation, we associated various "observable" quantities with various mathematical operations or "operators". As noted previously that while we are using these terms loosely at the moment, we will formally define them in the matrix mechanics section of this course where we introduce Dirac notation. From the correspondence principle, we associate each classical quantity with one of these operations and in particular note that $P \rightarrow -i\hbar\nabla$. We therefore write the Schrodinger equation (time independent) for a quantum rigid rotor as,

$$E\psi = \frac{L^2}{2mr^2}\psi \tag{43}$$

It is clear from this Hamiltonian that an equivalent problem is to solve the equation $\lambda\sigma = L^2\sigma$. Once we have the solution to this, we have the solution to the rigid rotor equation above. We can write the total angular momentum $L^2 \equiv L_x^2 + L_y^2 + L_z^2$, making the transformation to spherical coordinates ($(x, y, z) \rightarrow (r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)$) and transforming the ∇ function into spherical coordinates, we get the following equation in terms of θ and ϕ,

$$\begin{aligned} L^2\sigma &= \lambda\sigma \\ -\left(\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)\right)\sigma &= \lambda\sigma \end{aligned} \tag{44}$$

We can solve this equation through separation of variables where we write $\sigma(\theta, \phi) = \Theta(\theta)\Phi(\phi)$. This then leads to the following equation after some rearrangement

$$\frac{1}{\Phi}\frac{\partial^2}{\partial\phi^2}\Phi + \frac{\sin\theta}{\Theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)\Theta + \sin^2\theta\lambda = 0 \tag{45}$$

Notice the first term only depends on ϕ and is equal to terms independent of ϕ. For this to be valid for all ϕ this term must be equal to a constant. We therefore write,

$$\frac{d^2\Phi}{d^2\phi} = -a^2\Phi \tag{46}$$

The solution to this is $\Phi = \frac{1}{\sqrt{2\pi}}e^{ia\phi}$. For this function to be periodic in ϕ (that is to be single valued for example ϕ=0 and ϕ=2π are the same) a must be an integer $m = 0, \pm1, \pm2, ...$ Note that the factor $\sqrt{2\pi}$ is included in the solution of Φ to ensure that the function is normalized, that is that $\int d\phi\Phi^*\Phi = 1$. Substituting this solution for Φ back into our equation above, gives the following expression only in terms of θ.

$$\frac{-m^2}{\sin^2\theta} + \frac{1}{\Theta}\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)\Theta + \lambda = 0 \tag{47}$$

The next step is to solve the equation which depends on θ. In order to solve this, lets first introduce a substitution to a new independent variable $z = \cos(\theta)$ which varies over the range $z = [-1, 1]$. To replace $\Theta(\theta)$ by the function

$$P(z) = \Theta(\theta), \tag{48}$$

we need the following Jacobian

$$\frac{d\Theta}{d\theta} = \frac{dP}{dz}\frac{dz}{d\theta} = -\frac{dP}{dz}\sin\theta \tag{49}$$

and the fact that $\sin^2\theta = 1 - z^2$. Putting this all together gives the following new differential equation,

$$\frac{d}{dz}\left((1-z^2)\frac{dP(z)}{dz}\right) + \left(\beta - \frac{m^2}{1-z^2}\right)P(z) = 0 \tag{50}$$

We can solve this equation using the same method developed for the solution of the harmonic oscillator above. This is actually quite tedious in this case and the solution is outlined very nicely by Pauling and also in Eisberg used in the second year course. We will just note that the solution to this equation is the associated Legendre functions $P_l^m(\cos\theta)$:

$$P_l^m = (1-z^2)^{|m|/2}\frac{d^{|m|}}{dz^{|m|}}P_l(z) \tag{51}$$

where $P_l(z)$ are the ordinary Legendre polynomials. A general expression for these are given by Rodrigues' formula

$$P_l(x) = \frac{1}{2l!}\frac{d^l}{dx^l}(x^2-1)^l \tag{52}$$

The important thing to note here is that we have a second integer which enters in based on a recursion relation and reasoning similar to that in the harmonic oscillator. Recall that in the harmonic oscillator we required the series expansions to be finite at large x and ensured this by truncating them. If we were to go through the entire expansion here we would find a problem at $z = \pm 1$ requiring a similar truncation. The overall angular solutions to this problem are call spherical harmonics and have the form,

$$Y_l^m(\theta,\phi) \propto P_l^m(\cos\theta)e^{im\phi} \tag{53}$$

Out of this truncation in the series expansion, we have found a new integer l with $\lambda = l(l+1)$ where l and m are connected by $l = 0, 1, 2, 3...$ and $m = -l...+l$. We can therefore write the following solution to the original problem,

$$L^2\sigma = \lambda\sigma \rightarrow L^2Y_l^m = l(l+1)Y_l^m \tag{54}$$

To understand the physical significance of this, it is interesting to note what the object L_z is equal to. Using our concept of correspondence, this can be associated with the operation $\frac{-i\partial}{\partial\phi}$. Acting this on our angular solution gives a factor of m. We can then write the following equation,

$$L_zY_l^m = mY_l^m \tag{55}$$

The angular dependence in real space is plotted in Fig. 7. The ground state ($l = 0$) has no nodes in the angular dependence and as the energy increases the number of nodes increases. Three dimensional contours for $l = 0$ and $l = 1$ are shown in Fig. 8.
Therefore m is associated with the $z - axis$ projection of the total angular momentum.

6 Hydrogen atom (Morrison, Pauling-Chapter 5, Messiah-Chapter 9, Landau-Chapter 5)

To conclude the first section of course dealing with wave mechanics based on applications of Schrodinger equation, we discuss the solution to the Hydrogen atom. This is one of the few exactly solvable problems in quantum mechanics and of importance to quantum states in molecules. The underlying equation can be written as,

$$H = -\frac{\hbar^2}{2m}\nabla^2 + V(r) \tag{56}$$

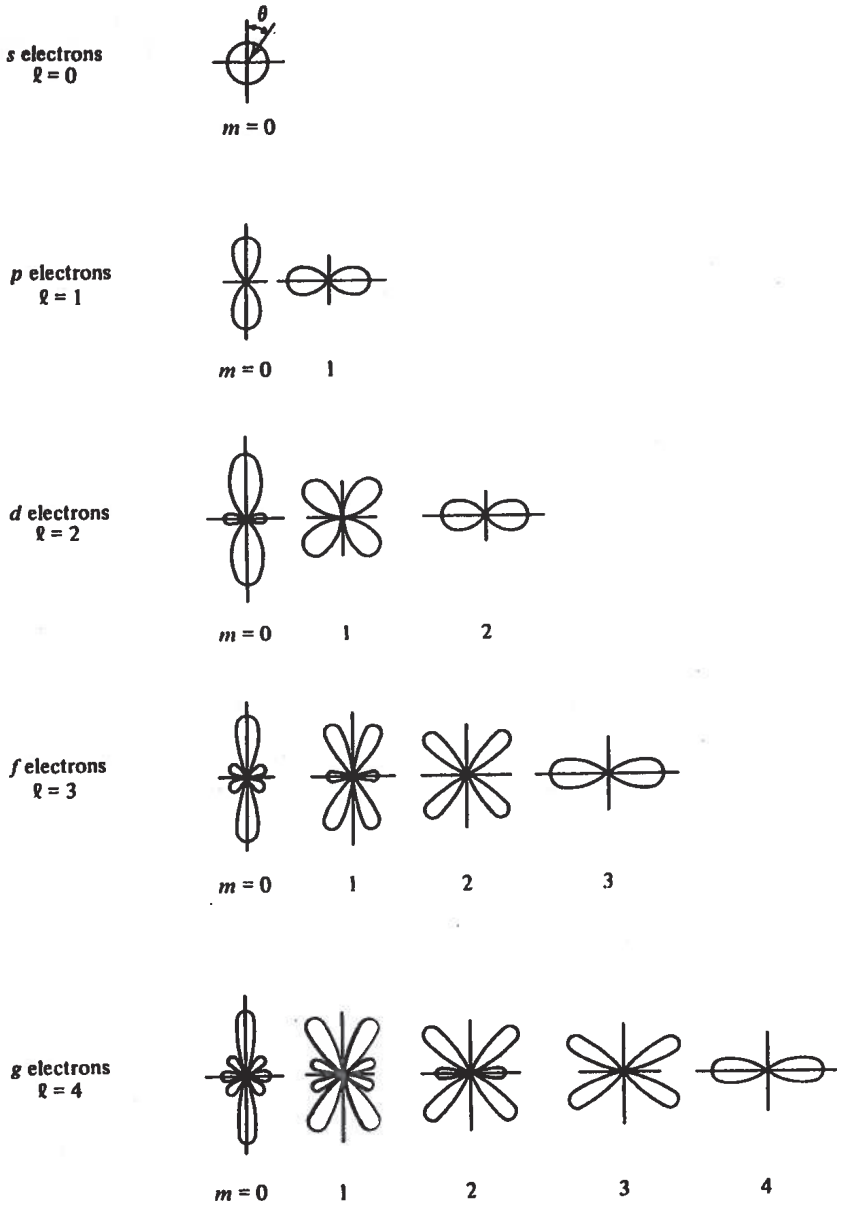

Figure 7: Figure taken Morrison, Estle, and Lane, *Quantum States of Atoms, Molecules, and Solids.*

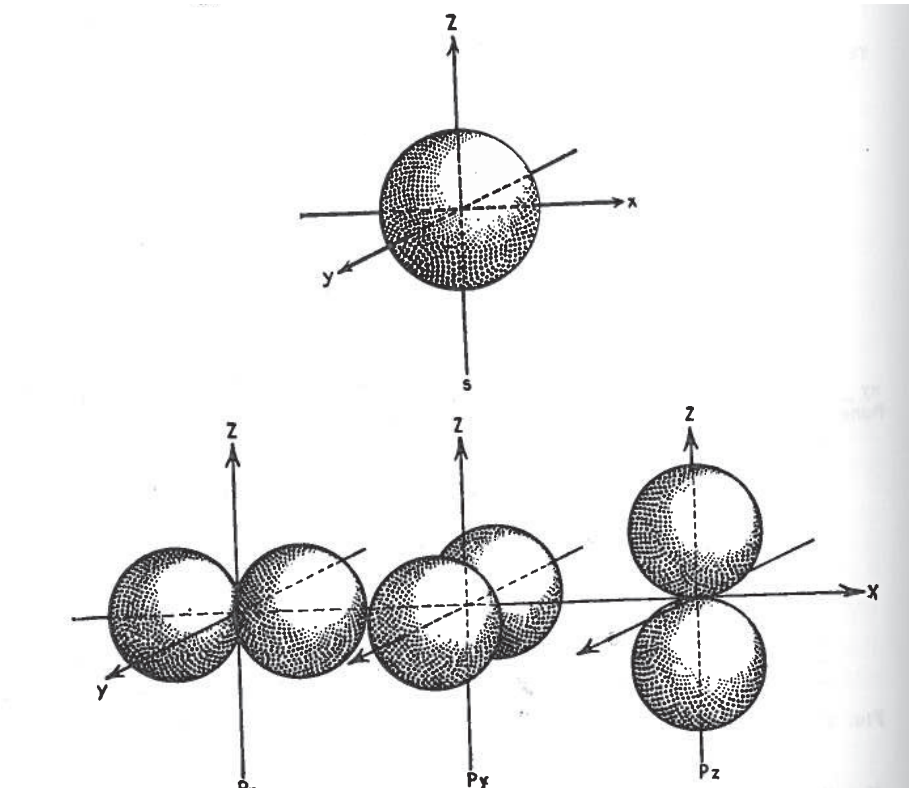

Figure 8: Figure taken Pauling, *Introduction to Quantum Mechanics.*

where the potential $V(r)$ takes the Coulomb form of $-\frac{e^2}{r}$ for the hydrogen atom which consists of a positive proton being circled by a negatively charged electron. For the moment we will only assume that $V(r)$ only depends on r and investigate solutions. There are several ways of solving this outlined in different books. All books use separation of variables and write separate equations for the angular and radial components. Pauling solves for three equations (r, ϕ, θ) while Messiah and Landau take advantage of the angular momentum operators discussed in the previous section. We will follow this later approach as it involves slightly less work and highlights the physical significance of the result. However, we will exploit Pauling's technique (used for the harmonic oscillator above) to solve the radial part of the equation.

Given the symmetry of the problem (note the equation does not explicitly depend on angles θ or ϕ), it is convenient to switch to spherical coordinates where,

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \tag{57}$$

the equation of motion then becomes,

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2} + \frac{2m}{\hbar^2}[E - V(r)] = 0 \tag{58}$$

If we introduce the operator L^2 we can reduce this to

$$\frac{\hbar^2}{2m}\left(-\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{L^2}{r^2}\right)\psi + V(r)\psi = E\psi \tag{59}$$

We know classically that angular momentum is conserved during motion in a central potential. We shall consider stationary states in which the angular moment l and the component m have definite

values. These values will then fix the angular dependence. We therefore will perform separation of variables to write ψ in terms of an angular and a radial part as follows.

$$\psi = R(r)Y_{lm}(\theta, \phi) \tag{60}$$

Since $L^2 Y_{l,m} = l(l+1)Y_{lm}$ we obtain for the radial function $R(r)$,

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) - \frac{l(l+1)}{r^2}R + \frac{2m}{\hbar^2}[E - V(r)]R = 0 \tag{61}$$

Let's solve the radial part and set $V(r) = -\frac{e^2}{r}$ and therefore specialize the problem to that of the hydrogen atom. Following Pauling, lets consider cases where the energy is insufficient to ionize the atom and hence consider negative energy values. It is helpful to group together the constants using the following substitutions,

$$\alpha^2 = -\frac{2mE}{\hbar^2} \tag{62}$$

$$\lambda = \frac{me^2}{\hbar\alpha} \tag{63}$$

and the new independent variable $\rho = 2\alpha r$. The wave equation then becomes,

$$\frac{1}{\rho^2}\frac{d}{d\rho}\left(\rho^2\frac{dS}{d\rho}\right) + \left(-\frac{1}{4} - \frac{l(l+1)}{\rho^2} + \frac{\lambda}{\rho}\right)S = 0 \tag{64}$$

where $0 \leq \rho < \infty$ and note that we have set $S(\rho) = R(r)$ and computed the appropriate Jacobian. Just like the case of the simple harmonic oscillator above, lets consider the case at large r when the potential is weak and the terms $\propto \frac{1}{r}$ are small. The equation then becomes,

$$\frac{d^2S}{d\rho^2} = \frac{1}{4}S \tag{65}$$

with solutions $S = e^{\frac{\pm\rho}{2}}$ of which only $S = e^{\frac{-\rho}{2}}$ is satisfactory based on the fact that wavefunction must be finite for all r. As in the simple harmonic oscillator, we now assume that the solution to this ODE has the form,

$$S(\rho) = e^{\frac{-\rho}{2}}F(\rho) \tag{66}$$

The equation satisfied by $F(\rho)$ can be derived to be,

$$\frac{d^2F}{d\rho^2} + \left(\frac{2}{\rho} - 1\right) + \left(\frac{\lambda}{\rho} - \frac{l(l+1)}{\rho^2} - \frac{1}{\rho}\right)F = 0 \tag{67}$$

where $0 \leq \rho < \infty$. The coefficients of $\frac{dF}{d\rho}$ and F possess singularities at the origin. To ensure that the solutions are well behaved and finite for all r, we make the additional substitution $F(\rho) = \rho^s L(\rho)$ in which $L(\rho)$ is the following power series which a non vanishing constant term.

$$L(\rho) = \sum_\nu a_\nu \rho^\nu \tag{68}$$

where $a_0 \neq 0$. Taking derivatives, we have the following,

$$\frac{dF(\rho)}{d\rho} = s\rho^{s-1}L + \rho\frac{dL}{d\rho} \tag{69}$$

and

$$\frac{d^2F(\rho)}{d\rho^2} = s(s-1)\rho^{s-2}L + 2s\rho^{s-1}\frac{dL}{d\rho} + \rho^s\frac{d^2L}{d\rho^2} \tag{70}$$

Putting this all together, our ODE for the radial part of the wavefunction becomes,

$$\begin{aligned}\rho^{s+2}\frac{d^2L}{d\rho} + 2s\rho^{s+1}\frac{dL}{d\rho} + s(s-1)\rho^sL + 2\rho^{s+1}\frac{dL}{d\rho} + 2s\rho^sL - ... \\ \rho^{s+2}\frac{dL}{d\rho} - s\rho^{s+1}L + (\lambda-1)\rho^{s+1}L - l(l+1)\rho^sL = 0\end{aligned} \tag{71}$$

By design, L begins with terms a_0 and therefore the coefficients of ρ^s must vanish and for this to happen $[s(s-1)+2s-l(l+1)]=0$. This gives us a relation for s implying that $s=+l$ or $-(l+1)$. Of the two solutions $-(l+1)$ does not give acceptable solutions and we can therefore write

$$F(\rho) = \rho^l L(\rho) \tag{72}$$

and further substitution gives

$$\rho\frac{d^2L}{d\rho^2} + [2(l+1)-\rho]\frac{dL}{d\rho} + (\lambda-l-1)L = 0 \tag{73}$$

We now explicitly introduce the power series presented above ($L(\rho) = \sum_\nu a_\nu\rho^\nu$) and again the coefficients must vanish individually. We then come up with the following recursion relations.

$$(\lambda-l-1)a_0 + 2(l+1)a_1 = 0 \tag{74}$$

$$(\lambda-l-2)a_1 + (4(l+1)+2)\,a_2 = 0 \tag{75}$$

$$(\lambda-l-3)a_2 + (6(l+1)+6)\,a_3 = 0 \tag{76}$$

or for general coefficients of ρ^ν

$$(\lambda-l-1-\nu)a_\nu + (2(\nu+1)(l+1)+\nu(\nu+1))\,a_{\nu+1} = 0 \tag{77}$$

It can be argued, like in the case of the harmonic oscillator, that for any values of λ and l the series whose coefficients are determined by this formula must be truncated. For very large values of ν, the successive terms of an infinite series are too large and do not converge. For the series to break off we need the following relation

$$(\lambda-l-1-n') = 0 \tag{78}$$

which implies $\lambda = n$ where $n = n'+l+1$. n' is called the radial quantum number and n the total quantum number. The allowed radial function can then be written (finally),

$$R(r) = e^{\frac{-\rho}{2}}\rho^l L(\rho) \tag{79}$$

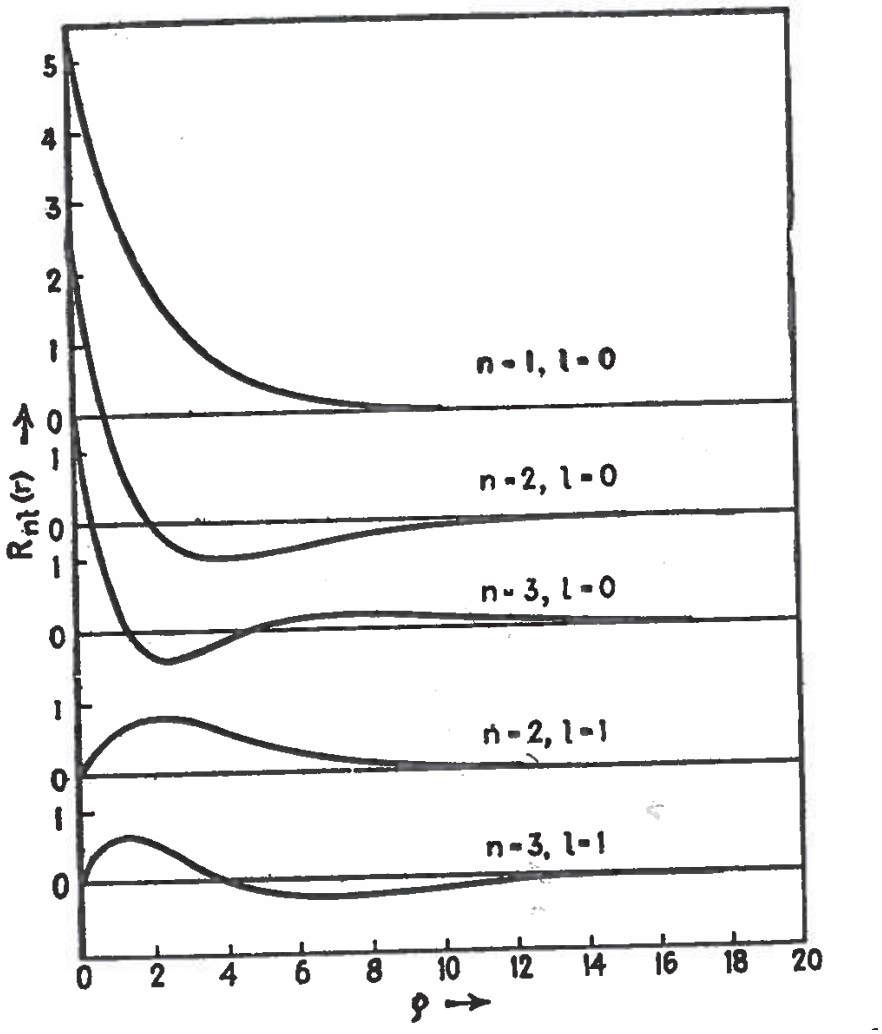

Figure 9: Figure taken Pauling, *Introduction to Quantum Mechanics.*

in which $L(\rho)$ is defined by the recursion formula above with $\lambda = n$. These functions are special functions and known as associated Laguerre functions.

Putting this all together in our expression earlier for λ gives the following for the energy eigenvalues (after dealing with the constants),

$$E_n = -13.6eV\frac{1}{n^2} \tag{80}$$

Except for $n = 1$, each energy level is degenerate being represented by more than one independent solution of the wave equation. If we introduce quantum numbers n, l, m as subscripts, the wave functions we have found as acceptable solutions of the wave equation are,

$$\psi_{nlm}(r, \theta, \phi) = R_{nl}(r)Y_{lm}(\theta, \phi) \tag{81}$$

The wavefunctions corresponding to distinct sets of values for n, l, m are independent with the allowed values being determined as follows,

$$\begin{aligned} m &= 0, \pm 1, \pm 2, ... \\ l &= |m|, |m| + 1, |m| + 2, ... \\ n &= l + 1, l + 2, l + 3, ... \end{aligned} \tag{82}$$

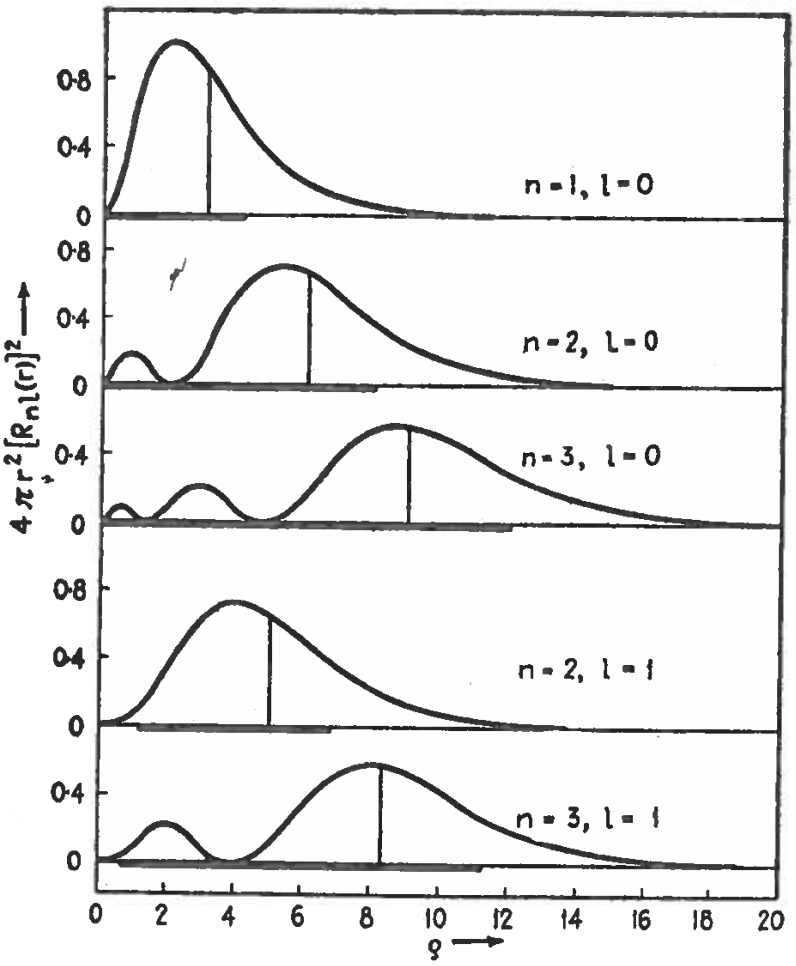

Figure 10: Figure taken Pauling, *Introduction to Quantum Mechanics.*

or

$$n = 1, 2, 3, ... \tag{83}$$

$$l = 0, 1, 2, ..., n-1$$

$$m = -l, -l+1, ... -1, 0, 1, ... + l - 1, +l \tag{84}$$

There are consequently $(2l+1)$ independent wave functions with given values of n and l.

There are are few qualitative things to note about the solutions. Figures 9 and 10 show plots of the radial wavefunctions as a function of ρ (which is proportional to r). Note that as n increases, the number of nodes increases just like in the harmonic oscillator and what was previously done with particle in the box. Note that the ground state has no nodes.

7 Formal presentation of wave mechanics

In the previous sections we have been studying applications of the Schrodinger equation to scattering, the harmonic oscillator, and the hydrogen atom. We took a probabilistic interpretation of the wave function $\psi(\vec{r}, t)$ and in particular noted that $|\psi(\vec{r}, t)|^2 d^3r$ represented the probablity of find at a time t the particle in a volume d^3r. From this application, we have gained some intuition on what defines a general wavefunction $\psi(\vec{r}, t)$. Here we summarize properties that we have found and generally must exist in real physical systems.

1) $\int d^3r |\psi(\vec{r}, t)|^2$=1: This makes sense as the probability cannot diverge and for the single particle problems discussed above, the particle must exist somewhere in space.

2) $\psi(\vec{r}, t)$ should be continuous and have continuous derivatives (except where there is a discontinuity in the potential V-remember the δ function potential).

Applying these "rules" helped us in solving the Dirac potential, the harmonic oscillator, the quantum rigid rotor, and finally the hydrogen atom.

7.1 A little bit of mathematics

Before developing a general theory of quantum mechanics, it is timely and important to review some linear algebra. Particularly, what defines a vector. In most early mathematics courses a vector is defined as something with magnitude and direction. This is completely wrong and is an unfortunate simplification.

7.2 Vector space

The previous section concludes the wave mechanics part of this course and we will now enter into the formal construction of quantum mechanics through the use of techniques from linear algebra. Given our experience with potential wells and the discussion in the previous sections, we now wish to formulate a general theory involving the set of square-integrable functions and have the property that they are defined everywhere and infinitely differentiable. Lets define $\mathcal{F}$ the set of wave functions composed of sufficiently regular functions that meet these criterion. While this is a slightly vague definition at the moment, we will make this more precise in the coming sections.

One can go through all of the axioms that define a vector space and formally prove that our properties (postulates above) derived from fairly simple physical arguments defines a vector space. For example, if two functions $\psi_1(\vec{r})$ and $\psi_2(\vec{r}) \in \mathcal{F}$, then $\psi(\vec{r}) = \lambda_1\psi_1(\vec{r}) + \lambda_2\psi_2(\vec{r})$ also $\in \mathcal{F}$ for λ_1 and λ_2 being two arbitrary complex numbers. Note that given the linear and homogeneous nature of the Schrodinger equation, it should be clear that ψ is a solution. We could also show that ψ is square integrable as required.

7.2.1 General definition of a Vector space

An unfortunate misconception is that a vector is an object that has direction and magnitude. This definition on its own is only true for a small subset of examples and a vector space is something much

more. I will cite the definition of a vector space below taken from L. Smith *Linear Algebra* (Springer, New York, 1998).

A vector space is a set $\mathcal{V}$, whose elements are called vectors, together with two operations. The first operation, called vector addition, assigns to each pair of vectors A and B a vector denoted as $A+B$, called their sum. The second operation, called scalar multiplication, assigns to each vector A and number r a vector denoted by rA. The two operatios are required to have the following properties:

AXIOM 1: $A+B=B+A$ for each pair of vectors A and B (commutative law of vector addition).
AXIOM 2: $(A+B)+C=A+(B+C)$ for each triple of vectors A, B, and C (associative law of vector addition).
AXIOM 3: There is a unique vector, called the zero vector, such that $A+0=A$ for every vector A.
AXIOM 4: For each vector there corresponds a unique vector $-A$ such that $A+(-A)=0$.
AXIOM 5: $r(A+B)=rA+rB$ for each pair of real number and r and each pair of vectors A and B.
AXIOM 6: $(r+s)(A)=rA+sA$ for each pair of real numbers, r and s, and each vector A.
AXIOM 7: $(rs)(A)=r(sA)$ for each pair of real numbers and r and s and each vector A.
AXIOM 8: For each vector A, $1A=A$.

Most people are familiar with vectors spaces defined in $\mathcal{R}^2$ (two dimensional cartesian space) and $\mathcal{R}^3$ (three dimensional cartesian space) which easily satisfy all of the axioms listed above, but it is interesting to think about what is NOT a vector space. For example, the set $(red, green, blue)$ is not a vector space as there is no $-green$. An interesting example of a vector space is the memory bits 0 and 1 and the operations associated with them. Finally, think about the properties of the functions $\in \mathcal{F}$ that are solutions of the Schrodinger equation (take for example the Hermite polynomial based solutions for the harmonic oscillator) and try to prove to yourself that they do satisfy the axioms listed above for a vector space.

7.3 Inner product

We will now define several general properties in vector spaces of interest. Lets define the scalar product for a pair ($\phi(\vec{r})$ and $\psi(\vec{r}) \in \mathcal{F}$. This operation is similar to the "dot product" that we are familiar with in vectors spaces $\in R^2$ and R^3. We can associate a complex number (denoted for now as (ϕ,ψ)) equal to.

$$(\phi,\psi) \equiv \int d^3r \phi^*(r)\psi(r) \tag{85}$$

Note that this integral always converges if ϕ and $\psi \in \mathcal{F}$. There are some fairly straightforward properties worth noting about this scalar product which is similar to the "dot" product you have definitely encountered previously.

$$(\phi,\psi) = (\psi,\phi)^* \tag{86}$$
$$(\phi,\lambda_1\psi_1+\lambda_2\psi_2) = \lambda_1(\phi,\psi)+\lambda_2(\phi,\psi) \tag{87}$$
$$(\lambda_1\phi_1+\lambda_2\phi_2,\psi) = \lambda_1^*(\phi_1,\psi)+\lambda_2^*(\phi_2,\psi) \tag{88}$$

If $(\phi,\psi)=0$, then $\phi(\vec{r})$ and $\psi(\vec{r})$ are said to be orthogonal. Examples of this can be taken from the previous sections. For example, in our study of the harmonic oscillator above we found a series of discrete energy levels described by the Hermite polynomials $H_n(x)$. For $H_n(x)$ and $H_m(x)$ it was shown in the exercises that the $\int_{-\infty}^{+\infty} H_n^* H_m dx$=0 (for $m \neq n$). These solutions are therefore orthogonal and $\in$ our new set of functions $\mathcal{F}$. Similar properties can be derived for the quantum rigid rotor (Y_l^m) and the hydrogen atom (Legendre polynomials etc.). As a final note, we also have the property that $(\psi,\psi)=\int d^3r|\psi(\vec{r})|^2$ is a real and positive number, and is zero only when $\psi(\vec{r})=0$.

7.4 Linear operators

A linear operator A is, by definition, "a mathematical entity which associate every function $\psi(\vec{r}) \in \mathcal{F}$ another function $\psi'(\vec{r})$, the correspondence being linear."

Formally,

$$\psi'(\vec{r}) = A\psi(\vec{r}) \tag{89}$$

and for two complex numbers λ_1 and λ_2,

$$A[\lambda_1\psi_1(\vec{r}) + \lambda_2\psi_2(\vec{r})] = \lambda_1 A\psi_1(\vec{r}) + \lambda_2 A\psi_2(\vec{r}) \tag{90}$$

Lets consider some simple examples of linear operators. For example the parity operator P behaves as follows $P\psi(\vec{r}) = \psi(-\vec{r})$. The position operator $X\psi(x,y,z) = x\psi(x,y,z)$, and the differential operator that differentiates with respect to x $D_x\psi(x,y,z) = \frac{\partial}{\partial x}\psi(x,y,z)$. Going back above we can find that our operators $L_{x,y,z}$ and $P_{x,y,z}$ we used for the quantum rotor and the hydrogen atom are also linear operators with similar definitions.

Consider the product of operators A and B, their product is defined by,

$$(AB)\psi(\vec{r}) = A[B\psi(\vec{r})] \tag{91}$$

In general, $AB \neq BA$ (!!!). Lets define the following commutator

$$[A, B] = AB - BA \tag{92}$$

Above, we talked a lot about X and P. Do the operators X and P actually commute? Lets compute the commutator acting on a wavefunction $\psi(x) \in \mathcal{F}$.

$$[X, P_x]\psi(x) = -i\hbar[x\frac{\partial}{\partial x} - \frac{\partial}{\partial x}x]\psi(x) = -i\hbar\left(x\frac{\partial}{\partial x}\psi(x) - \frac{\partial}{\partial x}(x\psi(x))\right) = i\hbar\psi(x) \tag{93}$$

based on the product rule of calculus. Since this is true $\forall\psi(x) \in \mathcal{F}$, then $[X, P_x] = i\hbar$. In the exercises, you will compute similar commutators for the angular momentum operators that we used in our study of the quantum rotor.

7.5 Discrete orthonormal basis in $\mathcal{F}$

We now discuss some properties and definitions of bases in $\mathcal{F}$. It is first important to quantify what we mean as orthonormal in general. Consider a countable set of functions of $\mathcal{F}$ labelled by a discrete index i. These could be our Hermite polynomials we found in the harmonic oscillator or the functions Y_l^m we found for the quantum rigid rotor. In general, consider the set $\{u_1, u_2, ...\} \in \mathcal{F}$. The set is *orthonormal* if

$$(u_i, u_j) = \int d^3r u_i^*(r)u_j(r) = \delta_{ij} \tag{94}$$

where δ_{ij} is defined as follows

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

The set of $\{u_1, u_2, ...\} \in \mathcal{F}$ constitute as *basis* if every function $\psi(\vec{r}) \in \mathcal{F}$ can be expanded in one and only one way in terms of the $u_i(\vec{r})$:

$$\psi(\vec{r}) = \sum_i c_i u_i(\vec{r}) \tag{95}$$

We now have the results that $(u_i, \psi) = \sum_j c_j (u_i, u_j) = \sum_j c_j \delta_{ij} = c_i$. Therefore, $c_i = (u_i, \psi) = \int d^3 r u_i^*(r)\psi(r)$. If we have two functions ϕ and $\psi \in \mathcal{F}$ - $\phi(\vec{r}) = \sum_i b_i u_i(\vec{r})$ and $\psi(\vec{r}) = \sum_j c_j u_j(\vec{r})$. Then $(\phi, \psi) = \sum_{i,j} b_i^* c_j \delta_{ij} = \sum_i b_i^* c_i$ and in particular $(\psi, \psi) = \sum_i |c_i|^2$. These relations are again generalizations of the familiar relations that we have worked with in vector spaces in R^2 and R^3.

7.6 The closure relation

The relation stated above - $(u_i, u_j) = \int d^3 r u_i^*(r) u_j(r) = \delta_{ij}$ implicitly implies that the set $\{u_i(\vec{r})\}$ are normalized to 1. Based on this, we can define the closure relation by noting

$$\begin{aligned} \psi(\vec{r}) = \sum_i c_i u_i(\vec{r}) = \sum_i (u_i, \psi) u_i(\vec{r}) = \sum_i \left(\int d^3 r' u_i^*(\vec{r'}) \psi(\vec{r'}) \right) u_i(\vec{r}) \\ = \int d^3 r' \psi(\vec{r'}) \left(\sum_i u_i^*(\vec{r'}) u_i(\vec{r}) \right) \end{aligned} \tag{96}$$

For this to be true, then the following relation must hold,

$$\sum_i u_i(\vec{r'}) u_i^*(\vec{r}) = \delta(\vec{r} - \vec{r'}) \tag{97}$$

This is known as the closure relation. If an orthonormal set satisfies the closure relation, it constitutes a basis.

7.7 Bases not belonging to $\mathcal{F}$

Above we discussed the properties of the set of functions $\{u_i(\vec{r})\}$ which are composed of square integrable functions. It can also be useful to introduce bases of functions not belonging to $\mathcal{F}$, but in terms of which any wavefunction $\psi(\vec{r})$ can be expanded. We will consider two examples.

7.7.1 Plane waves

The basis consider above, $\{u_i(\vec{r})\}$ consisted of square integrable functions. It is sometimes useful to consider bases that do not belong to this set, but that a wavefunction can be expressed. We have seen this in our analysis of the free particle Schrodinger equation where we found it helpful to consider functions which are not square integrable ($\sim e^{ikx}$). We found in the first workshop that we could construct sensible solutions which are square integrable by creating wave packets. You have seen in your Fourier analysis course the idea of Fourier transforms $\tilde{\psi}(p)$ of $\psi(x)$,

$$\begin{aligned} \psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp \tilde{\psi}(p) e^{ipx/\hbar} \\ \psi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx \psi(x) e^{-ipx/\hbar} \end{aligned} \tag{98}$$

For example, consider a wavefunction defined as $v_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar}$ - $v_p(x)$ is a plane wave with wave vector $p/\hbar$. You might recall that we used these plane wave solutions to motivate the Schrodinger equation in the first few sections of this class. Note that the integral of $|v_p(x)|^2$ diverges and therefore this function *cannot* belong to $\mathcal{F}$. Lets denote $\{v_p(x)\}$ as the set of all plane waves, that is, of all functions corresponding to the *continuous* index p (which varies over $-\infty$ to $+\infty$). We can write the following in analogy with our discussion of the set $\{u_i(\vec{r})\}$ given above.

$$\psi(x) = \int_{-\infty}^{\infty} dp \tilde{\psi}(p) v_p(x) \tag{99}$$

$$\tilde{\psi}(p) = (v_p, \psi) = \int_{-\infty}^{\infty} dx v_p^*(x) \psi(x) \tag{100}$$

therefore,

$$(\psi, \psi) = \int_{-\infty}^{\infty} dp |\tilde{\psi}(p)|^2 \tag{101}$$

This continuous basis also obeys closure

$$\int_{-\infty}^{\infty} v_p(x) v_p^*(x') = \frac{1}{2\pi\hbar} \int dp e^{ip(x-x')/\hbar} = \delta(x - x') \tag{102}$$

We now calculate the following inner product.

$$(v_p, v_{p'}) = \frac{1}{2\pi\hbar} \int dx e^{ix(p'-p)/\hbar} = \delta(p - p') \tag{103}$$

While the continuous and discrete basis (discussed earlier) may appear quite different, all of the relations carry over by making the following correspondence,

$$i \longleftrightarrow \vec{p} \tag{104}$$

$$\sum_i \longleftrightarrow \int d^3p \tag{105}$$

$$\delta_{ij} \longleftrightarrow \delta(\vec{p} - \vec{p}\,') \tag{106}$$

7.7.2 Delta function basis

As a second (and final) example of a continuous basis consider the set of functions of $\vec{r}$ $\{\xi_{\vec{r_0}}(\vec{r})\}$ labeled by the continuous index $\vec{r_0}$ defined by,

$$\xi_{\vec{r_0}}(\vec{r}) \equiv \delta(\vec{r} - \vec{r_0}) \tag{107}$$

We can show that these have similar closure relations as the continuous plane wave basis discussed above. In particular, all of our original relations for our original discussions of square integrable functions $\in \mathcal{F}$ with the following associations.

$$i \longleftrightarrow \vec{r} \tag{108}$$

$$\sum_i \longleftrightarrow \int d^3r_0 \tag{109}$$

$$\delta_{ij} \longleftrightarrow \delta(\vec{r_0} - \vec{r_0}\,') \tag{110}$$

These are similar to the association made above for the plane wave solutions. We will explore the properties of this example of a continuous basis if there is time, however we will make frequent use of the plane wave basis outlined above.

7.7.3 Generalization - Continuous orthonormal bases

Generalizing the discussion above, we can define a continuous orthonormal basis as a set of functions of $\vec{r}$, $w_\alpha(\vec{r})$, labeled by a continuous index α, which satisfy the orthonormalization

$$(w_\alpha, w_{\alpha'}) = \int d^3r \; w^*_\alpha(\vec{r}) w_{\alpha'}(\vec{r}) = \delta(\alpha - \alpha') \tag{111}$$

and closure

$$\int d\alpha \; w_\alpha(\vec{r}) w^*_\alpha(\vec{r}\,') = \delta(\vec{r} - \vec{r}\,') \tag{112}$$

Therefore, again we make the following correspondence which applies to a continuous orthonormal basis.

$$i \longleftrightarrow \alpha \tag{113}$$

$$\sum_i \longleftrightarrow \int d\alpha \tag{114}$$

$$\delta_{ij} \longleftrightarrow \delta(\alpha - \alpha') \tag{115}$$

7.8 Isn't this messy and annoying (yet)?

In the previous sections we used the physical fact that $\int d^3r |\psi(\vec{r},t)|^2$=1 to develop a wave vector formalism of quantum mechanics and defined some important mathematical operators (for bases $\in \mathcal{F}$ and not). We then discussed linear operators and some simple examples. Based on this, we talked about several basis states included discrete and continuous bases. We found that these could be based upon a number of different parameterizations and used the examples of spatial and momentum coordinates. This description of quantum mechanics is quite cumbersome and in general is problem specific (example real vs momentum space). In particular, the procedure of establishing a basis and then projecting it onto some coordinate system can be tedious (as evidence above in our discussion of the harmonic oscillator, the quantum rotor, and the hydrogen atom). It would really be nice if we could develop a mathematical framework which is arbitrary and applies to all of these basis states and also allow us to quickly and easily rederive the relations stated above for various basis states. This is what Dirac notation aims to do.

8 Simplifying life - Dirac notation and state space (Cohen-Tannoudji-Chapter 2, Messiah-Chapter 7)

While it is useful to project a our state onto a particular basis, it is rather cumbersome and problem specific. To get around this, we will introduce Dirac notation and the idea of a state space. We will start of by stating that each quantum state of a particle can be characterized by a state vector. We are going to now define the notation and rules in state space. In particular, lest consider a set of states in $\mathcal{E}$ which are a subspace of $\mathcal{F}$. Lets discuss the rules of vector calculation in $\mathcal{E}$.

8.1 Definition of Kets and Bras

As a matter of notation, any element or vector that lives in $\mathcal{E}$ is called a *ket vector*, or for the remainder of this course, a ket. It is represented by the symbol $|\rangle$. For example - $|\psi\rangle$. In particular, for any $\psi(\vec{r})$ $\in \mathcal{F}$, there is a $|\psi\rangle \in \mathcal{E}$.

We now define the scalar product. With each pair of kets $|\phi\rangle$ and $|\psi\rangle$, we associate a complex number which is their scalar product ($|\phi\rangle, |\psi\rangle$). To write this out in Dirac notation we need to introduce the concept of dual space and hence the "bra" vector. The scalar product is a linear function which

associates a complex number with every ket. It can be shown that the set of linear functions defined on the kets $|\psi\rangle \in \mathcal{E}$ constitues a vector space which is called the dual space of $\mathcal{E}$ and we will symbolize this by $\mathcal{E}^*$. Any element, or vector, of the space $\mathcal{E}^*$ is called a bra vector, or just a bra. We symbolize it as $\langle\ |$.

Note, linear functional and linear operator must not be confused (!!). In both cases, the operation is linear but a functional associates each ket with a complex number while the latter associates another ket. Associated with every ket in $\mathcal{E}$ there is a bra in $\mathcal{E}^*$. We define a linear function that associates (in a linear way) with each ket $|\psi\rangle \in \mathcal{E}$ a complex number which is equal to the scalar product $(|\phi\rangle, |\psi\rangle)$. Let $\langle\phi|$ be this linear functional defined as,

$$\langle\phi|\psi\rangle = (|\phi\rangle, |\psi\rangle) \tag{116}$$

Note that that the product is antilinear

$$(\lambda_1|\phi_1\rangle + \lambda_2|\phi_2\rangle, |\psi\rangle) = (\lambda_1^*\langle\phi_1| + \lambda_2^*\langle\phi_2|)|\psi\rangle \tag{117}$$

Therefore, the correspondence between kets and bras is analogous to the correspondence between wave functions of wave mechanics and their complex conjugate. We have now introduced two notations for the scalar product $(|\phi\rangle, |\psi\rangle)$ and $\langle\phi|\psi\rangle$. From now on, we will use the later notation ($\langle\phi|\psi\rangle$) and drop the former. For completeness, we now list (in Dirac notation) the properties of the scalar product.

$$\langle\phi|\psi\rangle = \langle\psi|\phi\rangle * \tag{118}$$

$$\langle\phi|\lambda_1\psi_1 + \lambda_2\psi_2\rangle = \lambda_1\langle\phi|\psi_1\rangle + \lambda_2\langle\phi|\psi_2\rangle \tag{119}$$

$$\langle\lambda_1\phi_1 + \lambda_2\phi_2|\psi\rangle = \lambda_1^*\langle\phi_1|\psi\rangle + \lambda_2^*\langle\phi_2|\psi\rangle \tag{120}$$

and we finally have $\langle\psi|\psi\rangle$ is real, positive, and zero if and only if $|\psi\rangle = 0$. Note that two kets $|\psi_1\rangle$ and $|\psi_2\rangle$ are orthogonal if $\langle\psi_1|\psi_2\rangle = 0$.

9 Linear operators

A linear operator associates with every $|u\rangle \in \mathcal{E}$, another ket $|v\rangle \in \mathcal{E}$ with the correspondence being linear,

$$\begin{aligned} A|u\rangle &= |v\rangle \\ A\left(\lambda_1|\psi_1\rangle + \lambda_2|\psi_2\rangle\right) &= \lambda_1 A|\psi_1\rangle + \lambda_2 A|\psi_2\rangle \end{aligned} \tag{121}$$

The product of two operators is defined as follows,

$$(AB)|\psi\rangle = A(B|\psi\rangle) \tag{122}$$

That is, B first operators on $|\psi\rangle$ to give the ket $B|\psi\rangle$, A then operates on this new ket $B|\psi\rangle$. These definitions are exactly analogous to the definitions in wave mechanics discussed above when we were consider wavefunctions $\psi(\vec{r})$. Like above, the commutator is defined as $[A, B] = AB - BA$. Recall that in general $AB \neq BA$.

It can be seen that the notation can be quite convenient. For example the inner product between $|\phi\rangle$ and $|\psi\rangle$ is simply $\langle\phi|\psi\rangle$ (noting the use of the dual space and one to one correspondence between kets and bras). There needs to be some caution associated with the notation as seen by considering the following,

$$|\psi\rangle\langle\phi| \tag{123}$$

This corresponds to an operator which can be seen by acting it on an arbitrary ket $|\chi\rangle$. The order of symbols is critical! Only complex numbers can be moved about without direct consequences. For example, $|\psi\rangle\lambda = \lambda|\psi\rangle$ and $\langle\psi|\lambda = \lambda\langle\psi|$. To see how this notation works in a bit more detail, lets consider the following "projector" operator

$$P_\psi = |\psi\rangle\langle\psi| \tag{124}$$

and apply it to an arbitrary ket $|\phi\rangle$

$$P_\psi|\phi\rangle = |\psi\rangle\langle\psi|\phi\rangle \tag{125}$$

Therefore, this operation gives a ket proportional to $|\psi\rangle$ with the coefficient of proportionality $\langle\psi|\phi\rangle$ being the scalar product between the two kets $|\psi\rangle$ and $|\phi\rangle$. Therefore, this operator projects onto a particular ket $|\psi\rangle$. This is confirmed by the fact that acting this operator twice is the same as once. For example, $P_\psi^2 = P_\psi P_\psi = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = P_\psi$. We can then project from one subspace of $\mathcal{E}$ to another through use of a generalization of this projector operator

$$P_q = \sum_{i=1}^{q} |\phi_i\rangle\langle\phi_i| \tag{126}$$

Acting this on an arbitrary $|\psi\rangle$ gives the linear superposition of the projection of $|\psi\rangle$ onto the various $|\phi_i\rangle$. We can use this to generalise the closure relation for a discrete basis as, $\sum_{i=1} |u_i\rangle\langle u_i| = 1$ where the sum is done over the entire subspace spanned by $|u_i\rangle$ and for a continuous basis as $\int d\alpha |u_\alpha\rangle\langle u_\alpha| = 1$. Here we have imposed the orthogonality relation $\langle u_i|u_j\rangle = \delta_{ij}$ and $\langle u_\alpha|u_{\alpha'}\rangle = \delta(\alpha - \alpha')$ for a continuous basis.

10 Hermitian conjugate

We have only discussed the operation of linear operators on kets $|\psi\rangle$. From the one to one correspondence between bras and kets, we can derive an analogous conjugation relation between linear operators. Let A be a linear operator and let $|v\rangle$ be the ket conjugate to the bra $\langle u|A$. As noted above, $|v\rangle$ depends antilinearly upon the bra $\langle u|$. This linear correspondence defines a linear operator which is termed Hermitian conjugate operator of A or sometimes called the adjoint operator of A and is symbolised by $A^\dagger$. Hence $|v\rangle = A^\dagger|u\rangle$.

An important property of Hermitian operators can be derived from noting that from the properties of scalar products $\langle u|v\rangle = \langle v|u\rangle^*$. Therefore,

$$\langle\psi|A^\dagger|\phi\rangle = \langle\phi|A|\psi\rangle^* \tag{127}$$

There is a number of properties of the Hermitian conjugate including $(A^\dagger)^\dagger = A$, $(\lambda A)^\dagger = \lambda^* A^\dagger$, and based on the linear nature of this operation $(A+B)^\dagger = A^\dagger + B^\dagger$. As a final note, $\langle A^\dagger\phi| = \langle\phi|(A^\dagger)^\dagger = \langle\phi|A$ and $\langle A^\dagger\phi|\psi\rangle = \langle\phi|A\psi\rangle$.

We now calculate $(AB)^\dagger$. Consider the ket $|\phi\rangle = AB|\psi\rangle \equiv A|\chi\rangle$, with $|\chi\rangle = B|\psi\rangle$. Therefore, $\langle\psi|(AB)^\dagger = \langle\chi|A^\dagger = \langle\psi|B^\dagger A^\dagger$. This gives us the result that $(AB)^\dagger = B^\dagger A^\dagger$. Note the reversal in the order of the operators in the expression. Moreover, the Hermitian conjugate of the operator $(|u\rangle\langle v|)^\dagger = |v\rangle\langle u|$. This can be proven by operating this on an arbitrary $|\phi\rangle$ and $|\psi\rangle$ - $\langle\psi|(|u\rangle\langle v|)^\dagger|\phi\rangle = [\langle\phi|(|u\rangle\langle v|)|\psi\rangle]^* = \langle\phi|u\rangle^*\langle v|\psi\rangle^* = \langle\psi|v\rangle\langle u|\phi\rangle$. Since this is true for $\forall$ $|\psi\rangle$ and $|\phi\rangle \in \mathcal{E}$ we have $(|u\rangle\langle v|)^\dagger = |v\rangle\langle u|$.

We can summarize the operation of taking the Hermitian conjugate as follows (from Messiah-Chapter 7): "replace everywhere the numbers by their complex conjugate, bras by the conjugate kets and vice versa, the operators by their Hermitian conjugates, and reverse in each term the order of the various symbols occurring there, that is to say the order of the bras, kets, and operators."

10.1 Hermitian operators

By definition, we refer to an operator as Hermitian if $A = A^\dagger$. Just to note (for terminology) that anti-Hermitian is defined as $I = -I^\dagger$. Note the product of two Hermitian operators is not necessarily Hermitean since $(HK)^\dagger = KH$. HK is only Hermitian if an only if H and K commute.

10.2 Unitary operators

An operator U is unitary if it is the inverse of its own Hermitian conjugate: $UU^\dagger = U^\dagger U = 1$. These will be the subject of the problem set.

11 Observables (Cohen-Tannoudji-Chapter 2, Messiah-Chapter 7)

11.1 Eigenvalues and vectors of an operator

We state that $|\psi\rangle$ is an eigenvector (or eigenket) of the linear operator A if

$$A|\psi\rangle = \lambda|\psi\rangle \tag{128}$$

where λ is a complex number. The eigenvalue λ is called nondegenerate when its corresponding eigenvector is unique to within a constant factor. However, if there exist a least two linearly independent kets which are eigenvectors of A with the same eigenvalue, this eigenvalue is said to be degenerate. For example, if λ is $g - fold$ degenerate, there correspond to it g independent kets $|\psi^i\rangle$ $(i = 1, 2, ...g)$ such that

$$A|\psi^i\rangle = \lambda|\psi^i\rangle \tag{129}$$

We can find the eigenvalues and eigenvectors from the characteristic equation $Det[A - \lambda I] = 0$.

There are several important properties of the eigenvalues of a Hermitian operator (that is when $A = A^\dagger$).

Property 1: The eigenvalues of a Hermitian operator are real. This follows if we take the definition of the eigenvalue problem above and therefore write $\langle\psi|A|\psi\rangle = \lambda\langle\psi|\psi\rangle$. We further note by the definition of a Hermitian operator, $\langle\psi|A|\psi\rangle^* = \langle\psi|A^\dagger|\psi\rangle = \langle\psi|A|\psi\rangle$. Since this is a real number and $\langle\psi|\psi\rangle$ is a real number, then λ, the eigenvalue, must be a real number.

Property 2: Two eigenvectors of a Hermitian operator corresponding to two different eigenvalues are orthogonal. This derives from considering two eigenvectors $|\psi\rangle$ and $|\phi\rangle$ of the Hermitian operator A.

$$A|\psi\rangle = \lambda|\psi\rangle \tag{130}$$
$$A|\phi\rangle = \mu|\phi\rangle \tag{131}$$

Since A is Hermitian, we can further note that $\langle\phi|A = \mu\langle\phi|$. Then multiplying the eigenvalue equations above, we obtain,

$$\langle\phi|A|\psi\rangle = \lambda\langle\phi|\psi\rangle \tag{132}$$
$$\langle\psi|A|\phi\rangle = \mu\langle\psi|\phi\rangle \tag{133}$$

Subtracting then gives the $(\mu - \lambda)\langle\phi|\psi\rangle = 0$. Since λ and μ are different eigenvalues by definition, then ϕ and ψ are orthogonal.

11.2 Definition

Consider a Hermitian operator A and kets $|u\rangle$ such that $A|u_k\rangle = a|u_k\rangle$ and therefore $\langle u_j|A = a'\langle u_j|$. Lets take the case that the subspace spanned by $\{u_i\}$ forms a complete orthonormal basis such that any ket $|\psi\rangle$ can be described as a linear combination of these orthonormal vectors. By definition, the Hermitian operator A is an observable if this orthonormal system of vectors forms a basis in the state space. This can be expressed formally through the closure relation.

$$\sum_{n=1}^{\infty} |u_i\rangle\langle u_i| = 1 \tag{134}$$

This mathematical relation applies to to a discrete basis. We can generalize it to the case of a continuous basis.

$$\int d\alpha |\alpha\rangle\langle\alpha| = 1 \tag{135}$$

The ultimate problem of knowing if a given Hermitian operator is an observable is in most cases a very difficult mathematical problem.

11.3 Commuting Observables (Cohen-Tannoudji-Chapter 2, Messiah - Chapter 8)

We now prove three useful theorems applying to commuting observables.

Theorem 1: If two operators A and B commute, and if $|\psi\rangle$ is an eigenvector of B, $B|\psi\rangle$ is also an eigenvector of A, with the same eigenvalue.

Proof: Take $A|\psi\rangle = a|\psi\rangle$ and apply B to both sides to give $B\left(A|\psi\rangle\right) = B\left(a|\psi\rangle\right)$. By definition A and B commute implying $A\left(B|\psi\rangle\right) = a\left(B|\psi\rangle\right)$. Therefore $B|\psi\rangle$ is an eigenvector of A with eigenvalue a.

Theorem 2: If two observables A and B commute, and if $|\psi_1\rangle$ and $|\psi_2\rangle$ are two eigenvectors of A with different eigenvalues, the matrix element $\langle\psi_1|B|\psi_2\rangle$ is zero.

Proof: $\langle\psi_1|(AB - BA)|\psi_2\rangle$=0 ; by the assumption that A and B commute. Therefore, using the eigenvalues we get the relation $(a_1 - a_2)\langle\psi_1|B|\psi_2\rangle = 0$. Given that $a_1 \neq a_2$, $\langle\psi_1|B|\psi_2\rangle = 0$.

Theorem 3: If two observables A and B commute, one can construct an orthonormal basis of the state space with eigenvectors common to A and B. The converse is also true that "if two observables A and B possess a common basis, they commute."

This follows from theorem 1 and theorem 2.

11.4 CSCO - Complete set of commuting observables

Consider an observable A and a basis of $\mathcal{E}$ composed of eigenvectors $|u\rangle$ of A. If none of the eigenvalues of A are degenerate, the various basis vectors of $\mathcal{E}$ can be labelled by the eigenvalue a_n. In other words, there exists only one basis of $\mathcal{E}$ formed by the eigenvectors of A. It is then said that the observable A constitutes, by itself, a complete set of commuting observales (CSCO). The key issue in any quantum mechanics problem, is to come up with a complete set of commuting observables (CSCO).

The concept of a CSCO is quite abstract. Lets apply this to the concept to the quantum rigid rotor discussed early on in this course where the Hamiltonian $H = \alpha L^2$. We found the eigenfunctions were the spherical harmonics which required two quantum numbers to uniquely identify them - l, m. We can then label the eigenstates as $|l, m\rangle$ using Dirac notation where $L^2|l, m\rangle = l(l+1)|l, m\rangle$ and

$L_z|l,m\rangle = m|l,m\rangle$. We note here that the operator L^2 by itself does not form a CSCO because there is a degeneracy associated with each eigenvalue l. The combination of L^2 and L_z does form a CSCO. The example of the harmonic oscillator we studied earlier is a bit more trivial as each eigenstate of H is non degenerate and therefor H forms a CSCO for the one dimensional harmonic oscillator we investigated early on in this course.

To define this more formally, "a set of observables A, B, C ... is called a complete set of commuting observables if 1) all of the observables A, B, C ... commute by pair; 2) specifying the eigenvalues of all the operators A, B, C ... determines a *unique* common eigenvector."

A central goal to every physical problem is the determination of a CSCO.

12 Postulates of Quantum Mechanics (Cohen-Tanoudji-Chapter 3)

First postulate: "At a fixed time t_0, the state of a physical system is defined by specifying a ket $|\psi(t_0)\rangle$ belonging to the state space $\mathcal{E}$."

Second postulate: "Every measureable physical quantity $\mathcal{A}$ is described by an operator A acting in $\mathcal{E}$; this operator is an observable."

Third postulate: "The only possible result of the measurement of a physical quantity $\mathcal{A}$ is one of the eigenvalues of the corresponding observable A."

Fourth postulate: "When the physical quantity A is measured on a system in the normlized state $|\psi\rangle$, the probability $P(a_n)$ of obtaining the eigenvalue a_n of the corresponding observable A is: $P(a_n) = \sum_{i=1}^{g_n} |\langle u_n^i|\psi\rangle|^2$, where g_n is the degree of degeneracy of a_n and $\{|u_n^i\rangle\}$" (i=1,2,...g_n) is an orthonormal set of vectors which forms a basis in the eigensubspace $\mathcal{E}_n$ associated with the eigenvalue a_n of A.

Fifth postulate: "If the measurement of the physical quantity A on the system in the state $|\psi\rangle$ gives the result a_n, the state of the system immediately after the measurement is the normalized projection $P_n|\psi\rangle/\sqrt{\langle\psi|P_n|\psi\rangle}$ of $|\psi\rangle$ onto the eigensubspace associated with a_n."

Sixth postulate: "The time evolution of the state vector $|\psi(t)\rangle$ is governed by the Schrodinger equation $i\hbar\frac{d}{dt}|\psi(t)\rangle = H(t)|\psi(t)\rangle$ where $H(t)$ is the observable associated with the total energy of the system."

13 Physical consequences of the postulates and our new theory

Before going to to apply this new theory of quantum mechanics to specific problems, we discuss some general physical consequences of the theory we have developed. We first discuss the notion of conserved quantities and then uncertainty relations.

13.1 Conserved quantities

In this section, we discuss an important physical consequence of the Schrodinger equation in terms of what formally defines a conserved quantity. Consider an arbitrary linear operator A, we can write the expectation value of this operator in general as follows,

$$\langle A\rangle(t) \equiv \langle\psi(t)|A(t)|\psi(t)\rangle \tag{136}$$

Differentiating and using the product rule of calculus

$$\frac{d}{dt}\langle\psi(t)|A|\psi(t)\rangle = \left(\frac{d}{dt}\langle\psi(t)|\right)A(t)|\psi(t)\rangle + \langle\psi(t)|A(t)\left(\frac{d}{dt}|\psi(t)\rangle\right) + ... \tag{137}$$

$$\langle\psi(t)|\frac{\partial A}{\partial t}|\psi(t)\rangle \tag{138}$$

Recall the Schrodinger equation above $i\hbar\frac{d}{dt}|\psi(t)\rangle = H(t)|\psi(t)\rangle$ gives the following,

$$i\hbar\frac{d}{dt}\langle A\rangle = \langle[A,H]\rangle + i\hbar\langle\frac{\partial A}{\partial t}\rangle \tag{139}$$

Therefore, a *constant of motion* is an observable A which does not depend explicitly on time and which commutes with H. Therefore, for A to be a constant of motion 1) $\frac{\partial A}{\partial t} = 0$ and 2) $[A,H]$=0.

Lets consider this in terms of the quantum rotor discussed above where $H = \alpha L^2$. Since L^2 commutes with the Hamiltonian and does not explicitly depend on time, it is a conserved quantity. Likewise, since H commutes with itself, energy is conserved quantity.

13.2 The uncertainty relations

Two observables, A and B are said to be compatible if they can be simultaneously determined. The precise basis of this derives from theorem derived above where if $|\psi\rangle$ is a eigenstate of B, then $B|\psi\rangle$ is an eigenstate of A. Note the converse is not (!!) true in general. If A and B do not commute, in general a state cannot be a simultaneous eigenvector of these two observables. The are said to be incompatible.

This leads to some important physical results regarding operators that do not commute. For example, this result seems to imply that measuring the position X, will introduce an uncertainty in the momentum P since these two operators do not commute. It is particularly interesting to think about the case of the quantum rigid rotor solved above. Recall that $[L^2, L_{x,y,z}] = 0$ implying that we can measure the total angular momentum and *only on* spatial projection. We cannot know, for example, both the eigenstates of L_z and L_x.

To quantify this, we will "derive" the position-momentum uncertainty relation. We will show that, in general, that if two observables A and B satisfy the equation,

$$[A,B] = i\hbar \tag{140}$$

then the product of their root-mean-square deviations will always remain greater than $\hbar^2/2$.

$$\Delta A\ \Delta B \geq \frac{\hbar}{2} \tag{141}$$

By definition,

$$\Delta A = (\langle A^2\rangle - \langle A\rangle^2)^{\frac{1}{2}} \tag{142}$$
$$\Delta B = (\langle B^2\rangle - \langle B\rangle^2)^{\frac{1}{2}}$$

Lets define the following observables

$$\tilde{A} = A - \langle A\rangle \tag{143}$$
$$\tilde{B} = B - \langle B\rangle \tag{144}$$
$$\tag{145}$$

Therefore,

$$\Delta A = \Delta \tilde{A} = \langle \tilde{A}^2 \rangle^{\frac{1}{2}} \tag{146}$$

$$\Delta B = \Delta \tilde{B} = \langle \tilde{B}^2 \rangle^{\frac{1}{2}} \tag{147}$$

Assume a dynamical state of the system is defined by the ket $|u\rangle$ normalized to unity. From Schwarz inequality to the vectors $\tilde{A}|u\rangle$ and $\tilde{B}|u\rangle$:

$$(\Delta \tilde{A})^2 (\Delta \tilde{B})^2 \equiv \langle u|\tilde{A}^2|u\rangle \langle u|\tilde{B}^2|u\rangle \geq |\langle u|\tilde{A}\tilde{B}|u\rangle|^2 \tag{148}$$

Separating $\tilde{A}\tilde{B}$ into Hermitean and anti Hermitian

$$\tilde{A}\tilde{B} = \frac{\tilde{A}\tilde{B} + \tilde{B}\tilde{A}}{2} + \frac{\tilde{A}\tilde{B} - \tilde{B}\tilde{A}}{2} = \frac{\tilde{A}\tilde{B} + \tilde{B}\tilde{A}}{2} + \frac{i\hbar}{2} \tag{149}$$

By direct substitution we can separate real and imaginary parts

$$\langle u|\tilde{A}\tilde{B}|u\rangle = \langle \frac{\tilde{A}\tilde{B} + \tilde{B}\tilde{A}}{2} \rangle + \frac{i\hbar}{2} \tag{150}$$

We now rewrite the Schwarz inequality

$$(\Delta \tilde{A})^2 (\Delta \tilde{B})^2 \geq \langle \frac{\tilde{A}\tilde{B} + \tilde{B}\tilde{A}}{2} \rangle^2 + \frac{\hbar^2}{4} \tag{151}$$

We therefore have the following relation we wanted to show,

$$\Delta A \ \Delta B \geq \frac{\hbar}{2} \tag{152}$$

There in the case of position and momentum where $[X, P] = i\hbar$, measuring one of these observables introduces and uncertainty in the other.

14 Ladder operators and the Harmonic oscillator (Messiah-Chapter 12,Cohen-Tannoudji-Chapter 5)

Previously we solved the Schrodinger equation for the harmonic oscillator in representation space. This was slightly cumbersome and involved the derivation of Hermite polynomials and recursion relations. Using Dirac notation we can solve this problem without projecting onto a particular basis (namely the position representation used previously) and solve the problem more generally and more directly by considering operators acting in state space. In this section we revisit this with operator techniques and introduce ladder (or creation and annihilation operators). We rewrite the Hamiltonian for quantum-mechanical problem of a particle of mass m in a one dimensional harmonic well.

$$H = \frac{P^2}{2m} + \frac{1}{2} m\omega^2 X^2 \tag{153}$$

Now, P and X are operators and are connected via the commutation relation $[X, P] = i\hbar$. The problem is of fundamental importance as it appears in the study of Quantum Electrodynamics and also Quantum Field theory. Before studying the solution to this Hamiltonian using operator techniques, we first make the following substitutions to remove constants and simplify the notation.

$$\hat{X} = \sqrt{\frac{m\omega}{\hbar}} X \tag{154}$$

$$\hat{P} = \frac{1}{\sqrt{m\hbar\omega}} P \tag{155}$$

With these substitutions the commutator $[\hat{X}, \hat{P}] = i$. The Hamiltonian can also be written as, $H = \hbar\omega\hat{H}$, with $\hat{H} = \frac{1}{2}\left(\hat{X}^2 + \hat{P}^2\right)$. We therefore seek solutions to the eigenvalue problem $\hat{H}|\psi_\nu\rangle = \epsilon_\nu|\psi_\nu\rangle$. If $\hat{X}$ and $\hat{P}$ were just numbers and this was a classical problem, we could write $\hat{X}^2 + \hat{P}^2$ as $(\hat{X} + i\hat{P})((\hat{X} - i\hat{P}))$, but $\hat{X}$ and $\hat{P}$ do not commute so these two expression are *not* equivalent. We can simplify the problem considerably by defining the following operators which mimic this step we would have done in more classical theory where these two variables commute.

$$a = \frac{1}{\sqrt{2}}\left(\hat{X} + i\hat{P}\right) \tag{156}$$

$$a^\dagger = \frac{1}{\sqrt{2}}\left(\hat{X} - i\hat{P}\right) \tag{157}$$

Note that while $\hat{X}$ and $\hat{P}$ are Hermitian, a and $a^\dagger$ are not, but are adjoints of each other. Noting that $[a, a^\dagger] = \frac{1}{2}[\hat{X} + i\hat{P}, \hat{X} - i\hat{P}] = \frac{i}{2}[\hat{P}, \hat{X}] - \frac{i}{2}[\hat{X}, \hat{P}]$, gives the following relation

$$[a, a^\dagger] = 1 \tag{158}$$

To see how this simplifies the notation we calculate $a^\dagger a$,

$$a^\dagger a = \frac{1}{2}\left(\hat{X} - i\hat{P}\right)\left(\hat{X} + i\hat{P}\right) = \frac{1}{2}\left(\hat{X}^2 + \hat{P}^2 + i\hat{X}\hat{P} - i\hat{P}\hat{X}\right) = \frac{1}{2}\left(\hat{X}^2 + \hat{P}^2 - 1\right) \tag{159}$$

Comparing this with our expression above for $\hat{H} = \frac{1}{2}\left(\hat{X}^2 + \hat{P}^2\right)$ gives the following,

$$\hat{H} = a^\dagger a + \frac{1}{2} \tag{160}$$

or

$$H = aa^\dagger - \frac{1}{2} \tag{161}$$

.
Note that the fact that $\hat{X}$ and $\hat{P}$ do not commute results in this extra zero point energy term of $\frac{1}{2}$. This key aspect of our new theory of quantum mechanics gives us this fundamentally different expression and new results like the Casimir effect discussed above.

Lets now define the operator N

$$N = a^\dagger a \tag{162}$$

The operator is Hermitian since $N^\dagger = a^\dagger \left(a^\dagger\right)^\dagger = a^\dagger a = N$. Moreover, we can then write

$$\hat{H} = N + \frac{1}{2} \tag{163}$$

therefore, the eigenvectors of N are the eigenvectors of $\hat{H}$ and vice versa. This may give the impression of an endless series of operator redefinitions to simplify $\hat{H}$, but to determine how this helps us with the goal of solving this Hamiltonian and determining the spectrum in state space, we need to derive a few properties of N, a, and $a^\dagger$. We first note that $[N, a] = -a$ and $[N, a^\dagger] = a^\dagger$. To solve for the spectrum of the harmonic oscillator in state space we therefore need to determine the eigenvectors and values of the operator N. We therefore seek solutions $N|\phi^i_\nu\rangle = \nu|\phi^i_\nu\rangle$. We consider a theorem outlined in Cohen-Tannoudji and in Messiah.

14.1 Properties of a and N

Theorem If $|\phi^i_\nu\rangle$ is an eigenvector of the operator N with the corresponding eigenavale ν, then the following are true.

Lemma 1: "The eigenvalues ν of the operator N are positive or zero."

PROOF: This can be shown by noting $|\left(a|\phi^i_\nu\rangle\right)|^2 = \langle\phi^i_\nu|a^\dagger a|\phi^i_\nu\rangle \geq 0$. We then note that $\langle\phi^i_\nu|a^\dagger a|\phi^i_\nu\rangle = \langle\phi^i_\nu|N|\phi^i_\nu\rangle = \nu\langle\phi^i_\nu|\phi^i_\nu\rangle$. Since $\langle\phi^i_\nu|\phi^i_\nu\rangle$ is positive it therefore follows that $\nu >\geq 0$. This is what we wanted to show.

Lemma 2: (properties of the vector $a|\phi^i_\nu\rangle$) "Let $|\phi^i_\nu\rangle$ be a non-zero eigenvector of N with eigenvalue ν (as stated above), we shall prove that (i) if $\nu = 0$, the ket $a|\phi^i_\nu\rangle$ is zero and (ii) if $\nu > 0$, the ket $a|\phi^i_\nu\rangle$ is a non-zero eigenvector of N with the eigenvalue $\nu - 1$."

PROOF: *(i)* Recall from above that $\langle\phi^i_\nu|a^\dagger a|\phi^i_\nu\rangle = \nu\langle\phi^i_\nu|\phi^i_\nu\rangle$, therefore the square of the norm of vector $a|\phi^i_\nu\rangle$ is zero if (and only if) this vector is zero. If $\nu = 0$ is an eigenvalue of N, all eigenvectors associated with this eigenvalue must satisfy the relation $a|\phi^i_\nu\rangle = 0$. This implies that $a^\dagger a|\phi^i_\nu\rangle = N|\phi^i_\nu\rangle = 0$. Therefore, any vector which satisfies $a|\phi^i_\nu\rangle = 0$ is an eigenvector of N with eigenvalue $\nu = 0$.

We have just worked out what the ground state is of N and hence $\hat{H}$. Note that ν must be zero or positive and this there implies that the ground state of $\hat{H} = N + \frac{1}{2}$ has a zero point ground state energy of $\frac{1}{2}$. Now that we have sorted out the ground state (and derived an interesting result pointing to the presence of a zero point energy), lets try and work out the excited state spectrum.

PROOF: (ii) Lets assume that $\nu > 0$ to begin to determine what the excited spectrum looks like. Applying what we know about commutators of a and $a^\dagger$ we get the following,

$$\begin{aligned}[N, a]|\phi^i_\nu\rangle = -a|\phi^i_{nu}\rangle = (Na - aN)\,|\phi^i_\nu\rangle = Na|\phi^i_\nu\rangle - aN|\phi^i_\nu\rangle = Na|\phi^i_\nu\rangle - \nu a|\phi^i_\nu\rangle \\ \Rightarrow -a|\phi^i_\nu\rangle = Na|\phi^i_\nu\rangle - \nu a|\phi^i_\nu\rangle\end{aligned} \tag{164}$$

Rearranging gives the following,

$$N\left(a|\phi^i_\nu\rangle\right) = (\nu - 1)\left(a|\phi^i_\nu\rangle\right) \tag{165}$$

Therefore $a|\phi^i_\nu\rangle$ is an eigenvector of N with eigenvalue $\nu - 1$. This is a nice relation and seems to imply a has the effect of lowering the energy, but we have not determined the spectrum or the details of our state space vectors $|\phi^i_\nu\rangle$. To carry on, lets determine some properties of $a^\dagger$.

Lemma 3: (properties of the vector $a^\dagger|\phi^i_\nu\rangle$) "Let $|\phi^i_\nu\rangle$ be a non-zero eigenvector of N with eigenvalue ν (as stated above), we shall prove that (i) the ket $a^\dagger|\phi^i_\nu\rangle$ is *nonzero* and (ii) the ket $a^\dagger|\phi^i_\nu\rangle$ is a eigenvector of N with the eigenvalue $\nu+1$."

PROOF: (i) Recall that $[a,a^\dagger]=1 \Rightarrow aa^\dagger - a^\dagger a = 1 \Rightarrow aa^\dagger = 1 + a^\dagger a$ (from the commutator relations discussed above). To prove this lemma we now note the following

$$\begin{aligned} |a^\dagger|\phi^i_\nu\rangle|^2 &= \langle\phi^i_\nu|a^\dagger a|\phi^i_\nu\rangle \\ &= \langle\phi^i_\nu|(1+a^\dagger a)|\phi^i_\nu\rangle = (1+\nu)\langle\phi^i_\nu|\phi^i_\nu\rangle \end{aligned} \tag{166}$$

By lemma 1 (proved above), ν is always postive or zero and therefore $a^\dagger|\phi^i_\nu\rangle$ must be always nonzero. Note, this distinguishes $a^\dagger|\phi^i_\nu\rangle$ from $a|\phi^i_\nu\rangle$ which can be zero. Lets now look at how $a^\dagger|\phi^i_\nu\rangle$ responds to the operator N.

Proof: (ii) The proof is similar and relies on the commutator $[N,a^\dagger]=a^\dagger$.

$$\begin{aligned} [N,a^\dagger]|\phi^i_\nu\rangle = a^\dagger|\phi^i_{nu}\rangle = \left(Na^\dagger - a^\dagger N\right)|\phi^i_\nu\rangle = Na^\dagger|\phi^i_\nu\rangle - a^\dagger N|\phi^i_\nu\rangle = Na^\dagger|\phi^i_\nu\rangle - \nu a^\dagger|\phi^i_\nu\rangle \\ \Rightarrow a^\dagger|\phi^i_\nu\rangle = Na^\dagger|\phi^i_\nu\rangle - \nu a^\dagger|\phi^i_\nu\rangle \end{aligned} \tag{167}$$

$$N\left(a^\dagger|\phi^i_\nu\rangle\right) = (\nu+1)\left(a^\dagger|\phi^i_\nu\rangle\right) \tag{168}$$

Therefore, the effect of $a^\dagger$ appears to be to raise the energy. Note, we have not assumed anything about ν or $|\phi^i_\nu\rangle$. We now have the necessary properties and tools to work out what the energy spectrum of N is and hence the hamiltonian $\hat{H}$.

14.2 Spectrum and basis of N and hence $\hat{H}$

Lets summarize what we just did in the previous section of formal operator algebra. We have re written the Hamiltonian in terms of the operator a, and $a^\dagger$. We then noted that since $\hat{H} = a^\dagger a + \frac{1}{2}$, that we might as well simplify this one more step and define $N = a^\dagger a$ and write $\hat{H} = N + \frac{1}{2}$. Since eigenvectors of N are also eigenvectors of $\hat{H}$, we then studied the properties of a and $a^\dagger$.

Given the theorem above applying to $a|\phi^i_\nu\rangle$ whose eigenvalue of N is $\nu-1$, we can apply this theorem to $a^2|\phi^i_\nu\rangle$, $a^3|\phi^i_\nu\rangle$, etc. The implication of the theorem is that the eigenvalue spectrum of these is $\nu-2$, $\nu-3$ etc. However, the set of eigenvalues has a lower limit (also from the theorem above) and that limit is 0. The implication of this fact is that ν must be an integer n.

Conversely, we can also apply this logic to $a^\dagger|\phi^i_\nu\rangle$ and derive the eigenvectors of repeated operations of $a^\dagger$ are $\nu+1$,$\nu+2$ etc. The integer nature of the eigenvalues is easier to see from the repeated action of $a^\dagger$. We know that the ground state (from the theorem above) has an eigenvalue of ν=0. Acting $a^\dagger$ again gives an eigenvalue of 1 and again 2 etc. This technique is actually quite general and will come up again particularly with angular momentum. In general, one can determine the ground state (or conversely the highest energy state) and through repeated operations of a or $a^\dagger$ obtain all of the eigenvalues and eigenvectors. This will be the topic of an exercise session.

In conclusion, the spectrum of eigenvalues of N is formed by the set of non-negative integers. Moreover, by repeated action of a and $a^\dagger$ on one of them, all of the eigenvectors can be obtained. Given the eigenvalues are integers, we now denote (for simplicity) the eigenvectors to be proportion to $|n\rangle = |0\rangle, |1\rangle, |2\rangle, |3\rangle....$ with eigenvalues of N equal to $0,1,2,3,....$ Given that a tends to lower the energy spectra it is referred to as the *annihilation* operator and since $a^\dagger$ has the effect of raising it is called the *creation* operator. We can work back our previous results in terms of Hermite polynomials by projecting these states onto the position representation $|x\rangle$. While this sounds rather abstract, it

is straightforward and is a topic of the exercises. We can therefore take our new basis states, which exist in state space, and project them onto any representation we like. This illustrates the elegance of using Dirac notation over our previous techniques of working in the position representation and solving rather difficult ODE's and PDE's.

Lets put these results together - given that $H = \hbar\omega\hat{H} = \hbar\omega\left(N + \frac{1}{2}\right)$, we can write the eigenvalue spectrum as

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right), n = 0, 1, 2, 3... \tag{169}$$

Note that since none of these eigenvectors are degenerate (also part of the exercises), the operator N (or H) form a CSCO. Therefore, once we have determined the properties of these eigenvectors and eigenvalues, we have formally solved this problem. While we have solved for the eigenvalue spectrum, we now must solve for the details of the eigenvectors. We will do this in the $|n\rangle$ representation.

The vector $|0\rangle$ associated with $n = 0$ is a vector that satisfies $a|0\rangle = 0$. According the theorem (namely lemma 3), the vector $|1\rangle$ corresponds to $n = 1$ and is proportional to $a^\dagger|0\rangle$:

$$|1\rangle = c_1 a^\dagger|0\rangle \tag{170}$$

where c_1 is some normalization constant which will now be determined. By imposing that the eigenvectors are normalized, we can determine what c_1 is as well as c_n.

$$\langle 1|1\rangle = |c_1|^2\langle 0|aa^\dagger|0\rangle = |c_1|^2\langle 0|1 + a^\dagger a|0\rangle \tag{171}$$

Since $|0\rangle$ is a normalized eigenvector, this implies that $|c_1|^2 = 1$ and that c_1=1. We can now proceed to get c_2

$$\langle 2|2\rangle = |c_2|^2\langle 1|aa^\dagger|1\rangle = |c_2|^2\langle 1|1 + a^\dagger a|1\rangle = 2c_2^2 \tag{172}$$

We therefore have,

$$|2\rangle = \frac{1}{\sqrt{2}}a^\dagger|1\rangle = \frac{1}{\sqrt{2}}\left(a^\dagger\right)^2|0\rangle \tag{173}$$

We can keep doing this for $|3\rangle$ etc.. and we would find that $c_n = \frac{1}{\sqrt{n}}$. We then write down the fully normalized eigenvectors of the harmonic oscillator in the $|n\rangle$ basis.

$$|n\rangle = \frac{1}{\sqrt{n!}}\left(a^\dagger\right)^n|0\rangle \tag{174}$$

Since H is Hermitian, the kets $|n\rangle$ corresponding to different values of n are orthogonal. Since each one is normalized as discussed above we have the following relation

$$\langle m|n\rangle = \delta_{m,n} \tag{175}$$

Since H is observable (by assumption) the set of $|n\rangle$ obey the following closure relation.

$$\sum_n |n\rangle\langle n| = 1 \tag{176}$$

Using the normalization above we also note for completeness the following properties of the creation and annihilation operators.

$$\begin{aligned} a|n\rangle &= \sqrt{n}|n-1\rangle \\ a^\dagger|n\rangle &= \sqrt{n+1}|n+1\rangle \end{aligned} \tag{177}$$

Note that the observables X and P can be written in terms of a and $a^\dagger$ and therefore we can switch between the number representation $|n\rangle$ and the position representation $|x\rangle$. This will be explored more in the exercises.

15 Angular momentum revisited - a general theory (Cohen-Tannoudji - Chapter 6, Messiah-Chapter 13)

We now revisit the angular momentum problem in quantum mechanics, but apply our use of raising and lowering operators to solve for the eigenvectors and values. Recall the fundamental formula for angular momentum,

$$\vec{L} = \vec{R} \times \vec{P} \tag{178}$$

Since we know the commutation relations for R and P observables, we can calculate the commutator relations for the operators L_x, L_y, and L_z. For example the commutator $[L_x, L_y]$:

$$[L_x, L_y] = [YP_z - ZP_y, ZP_x - XP_z] = [YP_z, ZP_x] + [ZP_y, XP_z] \tag{179}$$

and since YP_z commutes with XP_z, and ZP_y with ZP_x. We then have,

$$[L_x, L_y] = Y[P_z, Z]P_x + X[Z, P_z]P_y = -i\hbar YP_x + i\hbar XP_y = i\hbar L_z \tag{180}$$

We can repeat this for all possible iterations and we would get the following

$$[L_x, L_y] = i\hbar L_z \ , \ [L_y, L_z] = i\hbar L_x \ , \ [L_z, L_x] = i\hbar L_y \tag{181}$$

We have therefore established the commutation relations for the components of the angular momentum. We can adopt a more general point of view and define an angular momentum J as any set of three observables J_x, J_y, and J_z which satisfies,

$$[J_x, J_y] = i\hbar J_z \ , \ [J_y, J_z] = i\hbar J_x \ , \ [J_z, J_x] = i\hbar J_y \tag{182}$$

A rather compact way of writing this is in terms of the completely antisymmetrical tensor ϵ_{ijk} - $[J_i, J_k] = i\epsilon_{ijk}J^k$ (recall that repeated indices are summed over). We can also introduce the total angular momentum operator

$$J^2 = J_x^2 + J_y^2 + J_z^2 \tag{183}$$

which commutes with all three components of J

$$[J^2, J_i] = 0, i = x, y, z \tag{184}$$

This will be explicitly shown in the exercises using the same method as done above for the case of orbital angular momentum. We now seek out the eigenvalue spectra and eigenvectors of angular momentum. We have a problem though that the individual components do not commute and therefore we cannot have simultaneous eigenvectors. We will therefore seek eigenvectors and eigenvalues common to J^2 and J_z. To determine this, the procedure is actually quite similar to that of the harmonic oscillator outlined above.

It is convenient to introduce two new operators that are linear combinations of J_x and J_y.

$$\begin{aligned} J_+ &= J_x + iJ_y \\ J_- &= J_x - iJ_y \end{aligned} \tag{185}$$

Similar to the operators $a^\dagger$ and a of the harmonic oscillator, $J_\pm$ are not Hermitian but they are adjoints of each other. In the rest of this discussion we will only use the operators $J_\pm$, J_z, and J^2 and therefore we note the following relation,

$$J^2 = \frac{1}{2}(J_+J_- + J_-J_+) + J_z^2 \tag{186}$$

and therefore,

$$\begin{aligned} J_-J_+ &= J^2 - J_z(J_z + \hbar) \\ J_+J_- &= J^2 - J_z(J_z - \hbar). \end{aligned} \tag{187}$$

It is fairly straightforward to show the following commutator relations (based on the definitions presented above).

$$[J_z, J_+] = \hbar J_+ \ , \ [J_z, J_-] = -\hbar J_- \ , \ [J_+, J_-] = 2\hbar J_z \tag{188}$$

and

$$[J^2, J_z] = [J^2, J_-] = 0 \tag{189}$$

The eigenvalues of J^2 are positive and this can be seen by noting that it is the sum of positive definite Hermitian operators. Since J^2 commutes with each of the components of J, one can form a complete set of common eigenvectors of J^2 and one of its components, J_z for example. We will write the eigenvalues of J^2 as $j(j+1)$ from now on with $j \geq 0$.

Let $|jm\rangle$ be an eigenvector of J^2 and J_z with eigenvalues $j(j+1)\hbar^2$ and $m\hbar$ respectively. We therefore write,

$$J^2|j,m\rangle = j(j+1)\hbar^2|j,m\rangle \tag{190}$$

and

$$J_z|j,m\rangle = m\hbar|j,m\rangle. \tag{191}$$

Lets consider $J_\pm$, we now wish to find the properties of the operation $J_\pm|j,m\rangle$. As with the case of the ladder operators above, we first wish to consider $|J_+|j,m\rangle|^2$ and therefore we wish to understand terms which contain J_+J_- which we have written in terms of the operators J^2 and J_z above.
From the relations above we have the folowing,

$$\langle j,m|J_-J_+|j,m\rangle = \langle j,m|(J^2 - J_z(J_z+\hbar))|j,m\rangle = (j-m)(j+m+1)\hbar^2\langle j,m|j,m\rangle. \tag{192}$$

and

$$\langle j,m|J_+J_-|j,m\rangle = \langle j,m|(J^2 - J_z(J_z-\hbar))|j,m\rangle = (j+m)(j-m+1)\hbar^2\langle j,m|j,m\rangle. \tag{193}$$

Given that these cannot be negative, we get the following inequalities,

$$(j-m)(j+m+1) \geq 0 \tag{194}$$

and

$$(j+m)(j-m+1) \geq 0 \tag{195}$$

from which we can derive,

$$-j \leq m \leq j \tag{196}$$

Also, since the vanishing of a norm of a vector is a necessary condition for the vector vanishing, we have the following results,

$$J_+|j,m\rangle = 0 \tag{197}$$

when $(j-m)(j+m+1) = 0$. Also,

$$J_-|j,m\rangle = 0 \tag{198}$$

when $(j+m)(j-m+1) = 0$. Since m is in the interval $\pm j$ then these conditions amount to $m = +j$ and $m = -j$ respectively. For $m \neq j$, $J_+|j,m\rangle$ is a vector of angular momentum and hence,

$$J^2J_+|j,m\rangle = J_+J^2|j,m\rangle = j(j+1)\hbar^2J_+|j,m\rangle \tag{199}$$

We can then use the commutator $[J_z, J_+] = \hbar J_+$, to write the following.

$$J_zJ_+|j,m\rangle = J_+(J_z+\hbar)|j,m\rangle = (m+1)\hbar J_+|j,m\rangle \tag{200}$$

Therefore, $J_+|j,m\rangle$ is an eigenvector of J^2 and J_z with eigenvalues $j(j+1)\hbar^2$ and $(m+1)\hbar$ respectively. We can apply the same procedure to find $J_-|j,m\rangle$ is an eigenvector as well with eigenvalues $j(j+1)\hbar^2$ and $(m-1)\hbar$ respectively.
We can put all of these results together to formulate the following important theorem,

If $|j,m\rangle$ is a vector of angular momentum, then
(i) necessarily $-j \leq m \leq j$

(ii) if $m = j$, $J_+|j,m\rangle = 0$
if $m \neq j$, J_+ is necessarily a vector of angular momentum $(j, m+1)$ and a norm of $[j(j+1)-m(m+1)]$

(ii) if $m = -j$, $J_-|j,m\rangle = 0$
if $m \neq -j$, J_- *is necessarily a vector of angular momentum* $(j, m-1)$ *and a norm of* $[j(j+1)-m(m-1)]$

Therefore, for a fixed value of j, there exists an integer p such that $J_+^p|j,m\rangle$ is a non-zero eigenvector of angular momentum. Since $m + p = j$ (with p and integer), we have that $j - m$ is an integer. We can apply the same argument for $J_-^q|j,m\rangle$ implying that $p + q = 2j$ is also a non-negative integer. Since $2j$ is a positive integer, $j = 0$, $\frac{1}{2}$, 1, $\frac{3}{2}$, 2, ... The only only possible eigenvalues of J_z are integer and half integer numbers.

In a similar manner to the ladder operators above, we can construct the following relations,

$$J_+|j,m\rangle = \sqrt{j(j+1) - m(m+1)}\hbar|j,m+1\rangle. \tag{201}$$

and

$$J_-|j,m\rangle = \sqrt{j(j+1) - m(m-1)}\hbar|j,m-1\rangle. \tag{202}$$

16 Spin Operators (Landau - Chapter 5), Cohen-Tannoudji-Chapter 4, Messiah-Chapter 13)

In both classical and quantum mechanics, conservation of angular momentum is the direct result of of the isotropy of space. In the previous section, we discussed the matrix mechanics of the state $|l,m\rangle$ where $\langle\theta,\phi|l,m\rangle$ are the spherical harmonics Y_l^m discussed early on in the course with regards to the quantum rigid rotor. We have been discussing angular momentum of an object (for example an electron) circling around a fixed point. It turns out that there is also an internal angular momentum of a particle and this is demonstrated by the Stern-Gerlach experiment. Unfortunately, that the non-relativistic level, discussed here, this concept of an internal angular momentum needs to be added in "by hand" and does not come out (obviously) from theory. It does appear in a full relativistic treatment of the hydrogen atom and we will not attempt to pursue this here.

Experimentally, we are presented with the fact that there is an internal angular momentum with $j = \frac{1}{2}$. We call this new degree of freedom "spin" and outline some basic properties of this. We have seen above that with every measurable physical quantity there must be associated an observable. This is a Hermitian operator whose eigenvectors form a basis in state space. We will construct this state space and the properties of it.

16.0.1 Observable S_z

Given the doublet degeneracy associated with this internal angular momentum demonstrated by the Stern-Gerlach experiment, we know that our basis must consist of two states that will be denoted as $|+\rangle$ and $|-\rangle$. We will therefore define the following eigenvalues of S_z.

$$\begin{aligned} S_z|+\rangle &= \frac{1}{2}|+\rangle \\ S_z|-\rangle &= -\frac{1}{2}|-\rangle \end{aligned} \tag{203}$$

Invoking orthonormality for our new basis vectors,

$$\begin{aligned} \langle+|+\rangle = \langle-|-\rangle &= 1 \\ \langle+|-\rangle &= 0 \end{aligned} \tag{204}$$

Note that unlike above for the case of angular momentum where we required L^2 and L_z to form a CSCO, S_z alone forms a CSCO here. For completeness, we note the following closure relation,

$$|+\rangle\langle+| \; + \; |-\rangle\langle-| = 1 \tag{205}$$

We have therefore constructed a basis and properties of a two level system. Such formalism is important (particularly in quantum optics). We can construct a general vector in our new spin space as follows,

$$|\psi\rangle = \alpha|+\rangle + \beta|-\rangle \tag{206}$$

where

$$|\alpha|^2 + |\beta|^2 = 1 \tag{207}$$

In the $\{|+\rangle, |-\rangle\}$ basis, the matrix representing S_z is diagonal,

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

Like the case of angular momentum above, the three components of angular momentum do not commute. We can construct the operators S_x and S_y by first creating raising and lowering operators S_+ and S$_-$. These two operators can be created by inspection and also using the general theory created above.

$$S_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

and

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

From these two matrices and noting $S_x = \frac{1}{2}(S_+ + S_-)$ and $S_y = \frac{1}{(2i)}(S_+ - S_-)$, we get the following

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

and

$$S_y = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

These matrices following the commutator relations for angular momentum operators discussed above, namely $[S_y, S_z] = iS_x, [S_z, S_x] = iS_y, [S_x, S_y] = iS_z$. As a matter of notation, the three 2×2 matrices listed above are known as the Pauli spin matrices σ. We state the definitions here for completeness,

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

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

While this last step and definition might seem arbitrary (pointless) at this stage, these matrices have a number of helpful properties and become useful in field theory. For example, as outlined in Cohen-Tannoudji, they form a convenient basis of the 2×2 matrix space. They also can be used in dealing with two level atoms by assigning them an "fictitious" angular momentum (effectively just counting states).