PHYS 304 - Introduction to Quantum Mechanics

Jeff Young

November 8, 2021

Primary Reference:

Introduction to Quantum Mechanics, Third Edition, David J. Griffiths and Darrell F. Schroeter, Cambridge University Press 2018

The following Contents also serve as a syllabus.

Contents

Chapter 1

Introduction

Everything in this course is dealt with in the non-relativistic limit, which is implied throughout.

1.1   Why do we need quantum mechanics?

The vast majority of our direct interactions with the physical world, through the senses of touching, hearing, and seeing, can be *adequately* and quantiatively modelled using a combination of classical mechanics and classical electrodynamics. The underlying kinetic equation of classical mechanics (pertaining to the things we touch and our ears hear) is Newton's equation, and the kinetic equations defining the classical properties of the electric and magnetic fields associated with light (we see) are the Maxwell Equations. These two sets of equations are coupled via the Lorentz equation that describes the force exerted on particles in the presence of electric and magnetic fields.

The adjective *adequately* is used because the majority of everyday observed phenomena, and most engineering design, can be explained or accomplished using only these classical equations of motion, along with the underlying concepts and principles of classical mechanics and classical optics.

Although it is not usually emphasized, or the subtlties dwelled upon at the undergraduate level, each set of equations come in "microscopic" and "macroscopic" forms. Think about how one would have to go about rigorously solving the problem of how the trajectories of two billiard balls are alterred after a collision, using the classical equations of motion.....

Even if one was was ignorant of the atomic nature of matter, it is strictly necessary to break each billiard ball into a collection of differential mass elements, Δm_i, and solve

$\Delta m_i \frac{d^2 \Delta x_i}{dt^2} = F_i$, given a set of initial conditions $x_i(t = 0), \frac{dx_i}{dt}(t = 0)$, where $x_i(t)$ describes the position of the centre of mass of the differential mass element Δm_i, and F_i is the net force exerted on the ith particle. This of course leads to a massive number of coupled differential equations for all of the mass elements because F_i is determined by the forces exerted on the ith particle by all other mass elements in contact with it. This practically impossible mathmatical problem is rendered "trivial" to solve if one assumes that the macroscopic billiard ball is infinitely rigid and undergoes no deformation as it rolls and collides, which can only be possible if all of the *internal* forces between mass elements exactly cancel one another at all times. Then one gets the familar rigid body, or macroscopic form of Newton's equation, that relates the acceleration of the ball's *centre of mass position* to its total mass M, and the *external* force exerted on the ball as a whole. Similarly, to solve for the reflectivity of a metal mirror, strictly speaking one would have to use the Lorentz equation to solve for the effect of the electric field associated with the incident light wave on the motion of each electron in the metal surface of a mirror, and then solve the microscopic Maxwell Equations to figure out how each of those accelerating electrons radiates. In practice, by averaging all of those coupled microscopic equations of motion over macroscopic lengthscales, one obtains the macroscopic Maxwell Equations that describes the optical response of the metal mirror in terms of its complex dielectric response function, ϵ, and macroscopic current and charge *density distributions* that average over all the discrete electrons.

What messages should we take away from this reflection? That for the majority of everyday experiences, although real material objects are composed of a huge number of elementary particles, the details of how they are interconnected via highly complex *internal* forces is largely irrelevant to how they respond, as a well-defined collection, to macroscopic external forces, and this is because when averaged over many particles, these *internal* forces largely cancel each other. In practice, Newton's equation of motion very accurately describes the dynamics of the centre of mass of bodies composed of many elementary particles.

However, this does not mean that quantum mechanics is not playing a role in the problem: one needs to use quantum mechanics to understand how the material object forms in the first place!

1.2 Deconstructing the fundamentals of classical mechanics

From now on in this course, to introduce the concepts and tools of quantum mechanics, we will consider the dynamics of a single non-relativistic "particle of mass m" moving in some time-independent external potential V. Until the final section of the course, V is assumed to vary in just one dimension, labeled by position x. We will not consider the quantum mechanical treatment of things like light, that are classically treated as waves.

In this context, Newton's equation of motion is a second order differential equation for the temporal variation of the particle's position, which is equal to the local negative gradient of the external potential, divided by the particle mass. The classical solution to this problem is

to define $F(x) = -\frac{dV(x)}{dx}$, and solve $m\frac{d^2x}{dt^2} = F(x)$, which requires specification of the particle position and velocity at two specific times, usually taken to be the same time.

Thus, knowing what the position and momentum of the particle is at one instant in time, and knowing the potential landscape it experiences, Newton's equation provides a recipe for unambiguously predicting its future trajectory $(x(t))$ (and even recounting what its past trajectory must have been!).

Where, if anywhere, would uncertainty come into this problem? ....

From your laboratory courses, you should appreciate that any measurement has some associated uncertainty, be it due to the resolution of the instrument(s) used in the measurement, "noise", or practical limitations involved in using the apparatus, etc. Thus, the initial value of the position and velocity will never be precise, and one can truly only specify some range of possible trajectories consistent with the *measurement uncertainties* in the initial conditions. The "uncertainty" in the predicted trajectory has nothing to do with the equations of motion or their interpretation, but only to do with practical limitations imposed on specifying the initial conditions.

1.3 The recipe for solving the 1D particle motion problem using quantum mechanics

The solution of the 1D particle motion problem in quantum mechanics involves a different differential equation, where the function that is being solved for is **not** the particle's trajectory, but rather a *wavefunction* that *contains all the information it is possible to know about the particle's dynamics.* For instance, the particle's trajectory must somehow be evaluated from knowledge of its wavefunction. Now there are two ingredients to the recipe for solving the problem quantum mechanically; i) the differential equation that must be solved to determine the particle's wavefunction, and ii) a method of extracting information about its "trajectory" (more generally its dynamical evolution), once the wavefunction is known.

Obviously the quantum mechanical approach to solving the problem is both quantitatively and qualitatively completely different than the classical recipe: this course is devoted to familiarizing you with the underlying concepts and mathematical techniques involved in quantum mechanical approaches to solving particle kinetics problems.

The wavefunction that carries all of the information possible about the particle's dynamics is $\Psi(x,t)$, and the differential equation it satisfies is Schödinger's equation:

$$i\hbar\frac{\partial\Psi(x,t)}{\partial t} = \left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right)\Psi(x,t), \tag{1.1}$$

a partial differential equation in both position and time. The exact recipe for extracting information about measurable quantities from knowledge of the wavefunction at all x, t will be introduced in future lectures.

1.4   The wavefunction

The first, and arguably the most important conceptual step towards understanding the subtlies of this process is to appreciate how the wavefunction most directly connects to our classical notion of a particle. Noting that the wavefunction is in general a complex function (i.e. it in general has both real and imaginary parts at any x, t), the closest connection between the wavefunction and the classical notion of the corresponding particle is in the fact that the quantity,

$$\int_a^b |\Psi(x,t)|^2 dx \qquad (1.2)$$

is equal to the probability of finding the particle at some position between $x = a$ and $x = b$, at time t. Note that this statement is correct even if one assumes there is no measurement uncertainty in specifying the particle's position. Quantum mechanics has a built-in, intrinsic indeterminancy that only lets one specify the probability of measuring any particular value of some measurable quantity associated with the particle described by the wavefunction. Under some circumstances the wavefunction of a particle may be such that this intrinsic indetermancy is much less than the measurement uncertainty for at least a range of times, in which case the particle in that state would be behaving much like the solution of Newton's equation would predict, for that period of time, but this is **not** generally true.

One very important implication of this interpretation of the wavefunction has to do with the effect of the measurement process itself on the wavefunction. Without knowning any of the details of how the measurement of the particle's position is made (i.e. what time-dependent addtional terms are added to the potential it feels), one can deduce that if a measurement of the particle's position specifies the location of the particle at time t to be $x = c \pm \Delta x$, then immediately following the measurement, at time $t + \Delta t$, the squared magnitude of the wavefunction must be peaked at $x = c$ and have a limited range of Δx around $x = c$. Thus the process of a position measurement is said to "collapse the wavefunction" at the location the paricle is measured. Note that since the wavefunction is in general complex, this does not fully specify the wavefunction after the measurement, but it tells one alot! (One needs to know more details about how the measurement is carried out to extract any phase information).

Pause and summarize what has been covered in this first introduction to the amazing world of quantum mechanics;

1. although quantum mechanics, like classical mechanics, involves solving a differential equation to obtain the "answer", the entity that is being solved for, the particle's

wavefunction, is vastly different than the trajectory of the particle that is solved for in classical mechanics

2. the most direct way to connect a particle's wavefunction to its more familiar classical properties is by recognizing that in quantum mechanics,

$$\int_a^b |\Psi(x,t)|^2 dx \tag{1.3}$$

is equal to the probability of finding the particle at some position between $x = a$ and $x = b$, at time t.

3. since the complex wavefunction can in general spread out over a considerable distance at any given time, Eqn. 1.3 implies that there is an intrinsic indeterminancy associated with quantum mechanics: in general it is not possible to precisely predict measurements of the particle's trajectory, even if one knows the exact wavefunction

4. another implication of Eqn. 1.3 is that if one does measure its position at some time to be at a specific location, within some measurement uncertainty, immediately following the measurement the squared magnitude of the wavefunction must be localized around that position with a spread related to the measurement uncertainty. This is independent of the form of the wavefunction just before the measurement is performed.

1.5 exercises

1.6 Further insight into the interpretation of a particle dynamics through knowledge of its wavefunction

Another implication of Eqn. 1.3 is that

$$\int_{-\infty}^{+\infty} |\Psi(x,t)|^2 dx = 1 \tag{1.4}$$

which follows directly from the fact that it must be possible to find the particle somewhere.

Now lets reflect on what distinghishes a classical wave from a classical particle. The notion of a classical particle is based on it being at some well-defined (precisely *localized*) position at any given time. What is the purest form of a wave? Think of a pebble dropped in a calm pond, or the wake of a boat travelling through the water. In both cases, if one took a snapshot of the amplitude of the surface wave at some instant in time, it would look like a sinusoidal "wave" sketched in textbooks with an abscissa in units of position. Or, if one monitored the amplitude of the wave at a fixed location in space and plotted it as a function of time, the graph would again look like a "textbook" sinusoid, but now with the abscissa in units of time, rather than position. The former graph would have an

easily identifiable *wavelength*, and the latter an easily definable *period*. In practice, the amplitude of the sinusoidal disturbance graph grows and decays (is localized to some extent) in position or time. Already then, one can see that the conceptual distinction between waves and particles is somewhat vague. A "pure wave" goes on sinusoidally forever, is completely *delocalized*, and has a clear *wavelength* and *period*. A "pure particle" is infinitely localized, and has no discernable wavelength or period. Think now of a tsunami, which is practically and mathematically essentially a "solitary wave". While it is obviously a "wave", it is also highly localized and it would be difficult to ascribe to it either a period or a wavelength; so could it a particle?

To put this in a more formal, mathematical context, consider the 1D "pure" wave, $cos(kx - \omega t)$. This represents "a cosinusoidal wave traveling in the direction $sign(k/\omega)$, with wavelength $\lambda = 2\pi/|k|$, with a period $T = 2\pi/|\omega|$. The quantity k is referred to as a spatial frequency, or a wavevector. *As an aside, this is a good time to introduce complex waves, because remember that* $cos(kx) = \frac{1}{2}(e^{ikx} + e^{-ikx})$, *and therefore* $cos(kx)$ *is just equal to* $\mathscr{Re}\, e^{ikx}$. *Rather than having to work with cos and sin functions, it is much easier to work with the complex exponential forms, so lets consider our wavefunction to be* $e^{i(kx-\omega t)}$. Is this a valid wavefunction? One might be tempted to say "yes" because it is easily shown to be a valid solution of the Schrödinger equation) when $V(x) = 0$ (for specific values of k and ω). However the answer is technically "no" because it is not normalizable such that the integral of its magnitude over all x is unity. However, in reality you could imagine multiplying this infinitely extended "plane wave" with a gaussian envelope function, $\sqrt{\frac{1}{\sqrt{2\pi}\sigma}}e^{-\frac{1}{2}(\frac{x^2}{\sigma^2})}$ at $t = 0$, so that the resulting function is normalized, and is therefore a valid wavefunction. For very large σ (which can be arbitrarily large, so long as it is finite), one can effectively consider the wavefunction to be almost a pure, propagating plane wave.

What then is the probability distribution for measuring the particle described fully by the wavefunction,

$$\Psi(x, t) = \sqrt{\frac{1}{\sqrt{2\pi}\sigma}}e^{-\frac{1}{2}(\frac{x^2}{\sigma^2})}e^{i(kx-\omega t)}, \tag{1.5}$$

at $t = 0$? It is a Gaussian distribution spread with a standard deviation of σ about the origin. This particular solution of the Schrödinger equation therefore has the property that if one were to measure the position of the particle when it was in this state at $t = 0$, it could be found almost anywhere! More precisely, if one prepared the particle in this wavefunction many times in a row, and performed many measurements of the particle position, one on each "copy" of the wavefunction, the results of the measurements would form a histogram that in the limit of large measurement number, would take the form of the Gaussian with standard deviation σ. Clearly this solution of the particle's wavefunction bears no correspondence to a classical particle, (it is almost completely delocalized) which forces one to accept that in quantum mechanics, there are solutions for the particle's wavefunction that "behave" almost entirely wave-like.

Analyzing the above thought experiment, it is implicit that our notion of a classical

particle is based on it being at some well-defined (localized) location at any given time. How can any particle behave classically if it really should be described by quantum mechanics, which is fundamentally based on waves? Hint: see discussion of the tsunami above ...

The key to resolving this apparent mystery lies at the heart of the quantum mechanical *uncertainty principle* that will be encountered later in the course. To gain qualitative, empirical insight in regard to this particle-wave duality, imagine a large, long, rectangular wavepool. A linear wave-generating plunger spans the width of the pool and launches perfect "pure 1D waves" with period $T = 1/\nu$ when its motion is modulated sinusoidally with a frequency $2\pi\nu$. A photograph would show sinusoidal surface waves with a well-defined wavelength associated with ν and the speed of surface sound waves in the water. Changing ν would change λ in a perfectly predictable way, and the amplitude of the wave would be uniform away from the plunger.

If the action of the plunger was changed to include the superposition of two separate sinusoidal motions with (in general) different amplitudes and relative phases, one can easily see that the resulting wave motion would be much more complicated, and would depend, in general, on the relative amplitudes and phases. By adding many frequency components to the driving motion of the plunger, and carefully controlling their relative amplitudes and phases, it is actually possible to generate a "wave" that appears to be localized and "solitary" (the reader has to take this for granted at this point, but they should play with the online simulation tool to prove this). Thus a *superposition* of many "pure waves" can appear to behave much like a particle would behave (i.e. localized motion).

The mathematical underpinning of this crucial concept is based on the fact that k, x are conjugate variables (as are ω, t). Conjugate variables come in pairs that are Fourier Transform Duals of one another. Within factors of 2π to various powers, if you have a function of one real variable, say α and $f(\alpha)$, where $f(\alpha)$ is assumed integrable, then the Fourier transform of f is defined as,

$$F(\beta) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} f(\alpha)e^{-i\beta\alpha}d\alpha \tag{1.6}$$

where β is also real. Furthermore, $f(\alpha)$ can be expressed as an integral of $F(\beta)$,

$$f(\alpha) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} F(\beta)e^{i\alpha\beta}d\beta. \tag{1.7}$$

What this means is that if you have an arbitrary integrable function of position x (think some localized function that would resemble the location of a particle), it can be precisely

expressed as an infinite sum over a set of pure harmonic waves, e^{ikx}, with the exact nature of the pulse shape and location determined by a (unique) set of amplitudes and phases of each of the contributing waves, the $F(k)$. Later it will be shown more rigorously that there is a fundamental inverse relationship between the range of the wavelengths (or more precisely wavevectors k), Δk of "pure waves" with significant amplitudes that are needed to sum up to form a pulse with a given localization length, Δx. i.e. $\Delta k \propto \frac{1}{\Delta x}$. This can be explored using the Fourier analysis simulation tool.

1.7 How to calculate the expectation value of the velocity of the particle

Think of how one might formulate the problem of estimating the velocity of the particle, knowing only that the squared magnitude of the wavefunction, $\Psi(x,t)$, represents the probability distribution of where one would measure the particle to be at any given time, t. .....

To figure out a velocity one needs to calculate how the position of the particle changes between times t_1 and t_2. How does one define "the position of the particle" at any given time, quantum mechanically? ....

In the limit where $|\Psi(x,t_1)|^2$ and $|\Psi(x,t_2)|^2$ both represent clearly localized functions, with the centroid of their distributions shifted by more than their distribution widths between t_1 and t_2, then it should be clear that the velocity could be approximate as,

$$v\left(\frac{t_2 + t_1}{2}\right) \sim \left(\int_{-\infty}^{+\infty} x|\Psi(x,t_2)|^2 dx - \int_{-\infty}^{+\infty} x|\Psi(x,t_1)|^2 dx\right)/(t_2 - t_1) \qquad (1.8)$$

That is, to be pedantic, the velocity is equal to the time rate of change (derivative) of the *average position* one would expect to measure the particle in state $\Psi(x,t)$. Since any given set of measurements of the position of the particle at times t_2 and t_1 would yield different values of velocity, then this calculation must actually yield the *average* value of the velocity one would deduce by repeatedly measuring the wavefunction's position at two different times.

In quantum mechanics the average value of some measureable quantity is often referred to as the *expectation value*, and it is designated by using a pair of arrow brackets. Thus it seems reasonable to write

$$\langle v(t) \rangle = \frac{d}{dt}\langle x \rangle = \frac{d}{dt}\int_{-\infty}^{+\infty} x|\Psi(x,t)|^2 dx \tag{1.9}$$

Taking the time derivative inside the integral on the right hand side,

$$\langle v \rangle(t) = \int_{-\infty}^{+\infty} x\left(\left(\frac{\partial}{\partial t}\Psi^*(x,t)\right)\Psi(x,t) + \Psi^*(x,t)\frac{\partial}{\partial t}\Psi(x,t)\right)dx \tag{1.10}$$

By substituting for the time derivatives of the wavefunction using the Schödinger equation, one gets

$$\langle v \rangle(t) = \frac{-i\hbar}{2m}\int_{-\infty}^{+\infty} x\left(\left(\frac{\partial^2}{\partial x^2}\Psi^*(x,t)\right)\Psi(x,t) - \Psi^*(x,t)\frac{\partial^2}{\partial x^2}\Psi(x,t)\right)dx \tag{1.11}$$

Integrating twice by parts, and using the fact that the wavefunction must decay to zero at $x = \pm\infty$, one ends up with,

$$\langle v \rangle(t) = \frac{-i\hbar}{m}\int_{-\infty}^{+\infty} \Psi^*(x,t)\frac{\partial}{\partial x}\Psi(x,t)dx \tag{1.12}$$

Anticipating a more general result, we can suggestively rewrite the expressions for the expectation value of the position and velocity of the particle as

$$m\langle v \rangle(t) = \int_{-\infty}^{+\infty} \Psi^*(x,t)\left(-i\hbar\frac{\partial}{\partial x}\right)\Psi(x,t)dx \tag{1.13}$$

$$\langle x \rangle(t) = \int_{-\infty}^{+\infty} \Psi^*(x,t)x\Psi(x,t)dx \tag{1.14}$$

This then provides some insight regarding the second step in the recipe for solving quantum mechanical problems. To calculate the expectation value of any observable quantity, which for a particle moving in a 1D potential, can be expressed classically as a function of the particle's position and momentum, call this function $Q(x,p)$, one converts this to a *function of operators*, $\hat{Q}(\hat{x},\hat{p})$ representing the position and momentum operators, $\hat{x}$ and $\hat{p}$ respectively, and then evaluates,

$$\langle Q \rangle(t) = \int_{-\infty}^{+\infty} \Psi^*(x,t)\hat{Q}(\hat{x},\hat{p})\Psi(x,t)dx \tag{1.15}$$

From the above, it should be clear that $\hat{x} = x$, i.e. the operator associated with x is "multiply by x", while $\hat{p} = -i\hbar\frac{d}{dx}$, a much more subtle, differential operator; the momentum operator.

As an example, the expectation value of the total energy of the particle is

$$\langle E \rangle(t) = \int_{-\infty}^{+\infty} \Psi^*(x, t) \left(-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right) \Psi(x, t) dx \qquad (1.16)$$

Chapter 2

The Time-Independent Schrödinger Equation

2.1 A Powerful Solution Algorithm For Quantum Mechanical Problems Involving the Dynamics of a Single Non-relativistic Particle

The following outlines a process for breaking down the solution of the general Schrödinger equation (SE) into a sequence of steps that must be thoroughly understood in order to effectively set up and solve quantum mechanics problems. It is based purely on mathematical properties of partial differential equations like the SE. To fully appreciate the resulting algorithm that can be generally applied to find the solution of the SE, one starts by assuming that there are solutions that can be expressed as the product of one function that depends only on x, and another that depends only on t, i.e. assume $\Psi(x,t) = \psi(x)\phi(t)$. Substituting this into the SE, one gets

$$i\hbar \frac{1}{\phi(t)} \frac{d\phi(t)}{dt} = -\frac{\hbar^2}{2m} \frac{1}{\psi(x)} \frac{\partial^2 \psi(x)}{\partial x^2} + V(x). \tag{2.1}$$

For the equality to hold for all possible values of t and x, it must be that the two equations must be equal to some constant, call it E. Thus the single partial differential SE is replaced with two coupled differential equations, one, the *time-independent SE*,

$$\left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right) \psi(x) = E\psi(x), \tag{2.2}$$

and the second, *space-independent* SE,

$$i\hbar \frac{d\phi(t)}{dt} = E\phi(t) \tag{2.3}$$

For a given E, the solution of the space-independent SE can be taken simply to be

$$\phi(t) = e^{-i\frac{E}{\hbar}t} \tag{2.4}$$

by assuming that the normalization factor for the full wavefunction is absorbed in the solution of the time-indepedent SE. The time-independent SE is in general much harder to solve (depending on the form of $V(x)$), but one can still generally state that there will be a set of possible solutions, or eigen functions $\psi_n(x)$ for specific values of E labelled as E_n, of the eigen value problem defined by the equation,

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right)\psi_n(x) = E_n\psi_n(x) \tag{2.5}$$

Different potential landscapes defined by $V(x)$ will in general have very different sets of allowed eigen functions and eigen values, but it is this potential landscape, and the solution of the time-independent SE that determines the allowed values of E_n that are used in the corresponding solution of the full SE. Each $\Psi_n(x,t)$ defined by,

$$\Psi_n(x,t) = \psi_n(x)e^{-i\frac{E_n}{\hbar}t}, \tag{2.6}$$

where the $\psi_n(x)$ are assumed to be normalized, are referred to as *stationary state* solutions of the full SE for a given $V(x)$. While no proof will be given, it turns out that any wavefunction that satsifies the full SE can be expressed as a sum over all the system's stationary states, with constant expansion coefficients! This is an extremely powerful result, because it means that beyond solving for the eigen states of the time-independent SE, the only other step necessary to determine the full solution of the SE is to find the expansion coefficients! So, in equation form, any allowed solution of the full SE, $\Psi(x,t)$ can be expressed as

$$\Psi(x,t) = \sum_{n=1}^{\infty} c_n\psi_n(x)e^{-i\frac{E_n}{\hbar}t} = \sum_{n=1}^{\infty} c_n\Psi_n(x,t), \tag{2.7}$$

where each c_n is a constant. So, how does one find the expansion coefficients, c_n? Recall in classical mechanics one solves for the general form of the solution of Newton's equation for a given driving force function, which will have two arbitrary constants. To find these two constants and complete the solution it is necessary to specify the initial conditions (the position and the velocity of the particle at some time, call it $t = 0$). In analogy, in quantum mechanics, to find the exapansion coefficients c_n one has to specify the spatial dependence of the wavefunction at some time, call it $t = 0$. This leads to the equation

$$\Psi(x,0) = \sum_{n=1}^{\infty} c_n\psi_n(x) \tag{2.8}$$

The method by which one can use this to find all of the c_n will be covered later, but it can be qualitatively understood by thinking of the function $\Psi(x,0)$ as some arbitrary vector in a high dimensional vector space, and each of the $\psi(x)$ functions as an orthonormal set of

unit vectors that completely span that high dimensional vector space. The determination of the c_n then basically involves projecting $\Psi(x,0)$ onto each of these basis vectors $\psi(x)$.

It is now possible to summarize the key steps involved in solving any quantum mechanics problem (for particles moving in a 1D, time-independent potential);

1. Solve the time-independent Schrödinger equation,

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right)\psi(x) = E\psi(x), \tag{2.9}$$

   mathematically an eigen value problem, for the set of eigen functions of position $\psi_n(x)$, and the corresponding eigen values E_n,

2. Use the particle's wavefunction at time $t = 0$ to evaluate the constant expansion coefficients c_n that uniquely specify how much of each eigen function $\psi_n(x)$ is needed to make up $\Psi(x,0)$,

$$\Psi(x,0) = \sum_{n=1}^{\infty} c_n \psi_n(x), \tag{2.10}$$

3. Then the complete solution of the problem is given by

$$\Psi(x,t) = \sum_{n=1}^{\infty} c_n \psi_n(x) e^{-i\frac{E_n}{\hbar}t}. \tag{2.11}$$

Before going into specific examples of how to solve the eigen value problem (the time-independent SE), and do the projection operations to find the c_n, first explore some general properties of these eigen states, and the corresponding stationary states.

Begin by finding the expectation value of any observable quantity, $Q(x,p)$ if the particle is "in a stationary state". The phrase "in a stationary state" means that the particle's dynamics is fully described by one and only one of the stationary states labelled by E_n, with the associated wavefunction $\psi_n(x)e^{-i\frac{E_n}{\hbar}t}$.

The expectation value is given by

$$<Q>(t) = \int_{-\infty}^{+\infty} \psi_n^*(x)e^{i\frac{E_n}{\hbar}t}\hat{Q}(\hat{x},\hat{p})\psi_n(x)e^{-i\frac{E_n}{\hbar}t}dx \tag{2.12}$$

but the operators $\hat{x}$ and $\hat{p}$ do not "operate on" (i.e. they "pass through") the exponential functions that depend only on E_n and t, so

$$<Q>(t) = \int_{-\infty}^{+\infty} \psi_n^*(x)\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right)\psi_n(x)dx, \tag{2.13}$$

which is independent of time. This is why the $\Psi_n(x,t)$ are referred to as "stationary states".

Another interesting property of the stationary states is revealed by evaluating the expectation value of the total energy of the particle in an arbitrary stationary state.

$$< E > (t) = \int_{-\infty}^{+\infty} \psi_n^*(x) \left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right) \psi_n(x) dx, \tag{2.14}$$

but we know that

$$\left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right) \psi_n(x) = E_n \psi_n(x) \tag{2.15}$$

since this is the defining eigen value problem for $\psi_n(x)$, so it follows that $< E >= E_n$ when the particle is assumed to be in the nth stationary state. This should explain why the symbol E was chosen for the constant when first deriving the stationary states from separation of variables!

Evaluating the expectation value of the total energy squared trivially yields $< E^2 >= E_n^2$ by repeated application of the eigen value equation above. Why is this interesting? Because more generally the result of a measurement of the total energy of a particle that is described by some arbitrary wavefunction $\Psi(x,t)$ will have a range of possible outcomes (just like the measurement of its position is in general given by some probability distribution $|\Psi(x,t)|^2$. It should be clear that a good measure of the range of energies that might be measured is the expectation value of the variance

$$< (E- < E >)^2 > (t) =< E^2 > (t)- < E >^2 (t) \tag{2.16}$$

But from above, this is zero for all times if the particle is in a stationary state, since then $< E^2 >= E_n^2$ and $< E >= E_n$! Thus one can conclude that the result of measuring the total energy of the particle when it is in a stationary state is precisely E_n, with no variance.

Within the scope of particles restricted to move in 1D, one problem differs from another only in the differences in $V(x)$, so the next sections will provide examples of how to use the above recipe for a few different potential profiles.

2.2 Example 1: Finding the Stationary States for $V(x)$ Corresponding to an Infinite Square Well Potential

This is a particularly simple example to start with, where $V(x)$ is infinite for $x < 0$ and $x > a$, and $V(x) = 0$ for $0 < x < a$. If one thinks classically, a region of infinite potential represents a space where the particle cannot enter (it can't have the infinite kinetic energy needed to overcome, or make it up the potential barrier). Thus we can assume that $|\Psi(x,t)|^2 = 0$ for $x < 0$ and $x > a$, which means that the wavefunction itself must also disappear in those regions. That was simple! The solution of the time-independent SE then only has to be done explicitly in $0 < x < a$, where $V(x) = 0$, so the eigen value problem becomes:

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) = E\psi(x) \tag{2.17}$$

Clearly any function $Ae^{ikx} + Be^{-ikx}$ with A and B arbitrary constants, satisfies this differential equation in $0 < x < a$, and is therefore a valid candidate eigen function in that region, with an eigen value $E(k) = \frac{\hbar^2}{2m}k^2$. But to satisfy the full time-independent SE for the true $V(x)$ over the entire x axis, the function has to vanish at $x = 0$ and $x = a$. It is this requirement that limits the allowed eigen values and eigen functions, since this is only possible if $k = \pm n\pi/a$, with n a positive integer, and $B = -A$. The true eigen functions of the full potential are therefore

$$\psi_n(x) = A\sin(\frac{n\pi}{a}x) \tag{2.18}$$

for n=1,2,.... (any nonzero positive integer). The constant A is determined from the normalization condition, $\int_{-\infty}^{+\infty} \Psi_n^*(x,t)\Psi_n(x,t)dx = 1$ so it should be clear that $A = \left(\frac{2}{a}\right)^{\frac{1}{2}}$. The stationary states are therefore

$$\Psi_n(x,t) = \left(\frac{2}{a}\right)^{\frac{1}{2}} \sin(\frac{n\pi}{a}x)e^{-i\frac{\hbar n^2\pi^2}{2ma^2}t}. \tag{2.19}$$

This completes the full solution of one of the simplest quantum mechanical problems!

2.2.1 Reflect on some of the properties of the solution

A few point-form observations that are quantitatively specific to this example, but have quite general applicability at a qualitative level:

1. There are an infinite number of eigen functions and corresponding eigen energies, and therefore there are an infinite number of stationary states.

2. Each of the eigen states, labelled by n, is characterized by the number of nodes of the wavefunction, which are n-1 in number, and therefore the expected value of the energy of each stationary state increases monotonically with the number of nodes in the associated wavefunction.

3. All of the eigen energies are positive.

4. The separation in energy between any two eigen states labelled by n and $m \neq n$ increases as the width of the well decreases, as $\frac{1}{a^2}$. Related, the lowest energy state for the particle in this potential well has an energy that increases as $\frac{1}{a^2}$. This is an example of a very general quantum effect referred to as 'size quantization'.

2.3 Evaluating the expansion coefficients c_n of a general wavefunction

The only thing left to complete the quantum mechanics problem-solving-recipe is to explain how to obtain the all-important expansion coefficients, c_n. The c_n can be obtained using the

fact that the eigen functions of the time-indepedent SE are known to form what is called an *orthonormal set* of functions, which means that any arbitrary, well-behaved function on $-\infty < x < +\infty$ can be expanded uniquely as a superposition of all of the eigen states, with expansion coefficients that are complex numbers. The orthogonality property, or the *inner product* of the normalized $\psi_n(x)$ is defined by

$$\int_{-\infty}^{\infty} \psi_n^*(x)\psi_m(x)dx = \delta_{n,m} \tag{2.20}$$

where δ_n, m is the Kronecher delta function that is unity when $n = m$, and zero otherwise. Since

$$\Psi(x,0) = \sum_{n=1}^{\infty} c_n\psi(x), \tag{2.21}$$

if one "operates" on both sides of this equation with the integral operator

$$\int_{-\infty}^{\infty} dx\psi_m^*(x), \tag{2.22}$$

then immediately one finds that

$$c_m = \int_{-\infty}^{\infty} \psi_m^*(x)\Psi(x,0)dx \tag{2.23}$$

The fact that (not proved here) *any* valid function of x compatible with the boundary conditions of the problem can be uniquely expanded in terms of the set of eigen states of the time-independent SE means that this set of functions are both orthonormal, and *complete.*

2.4 The Complete Quantum Mechanical Problem Solving Recipe

Although the following is largely a repetition of material previously summarized, this recipe is so central to any introduction to quantum mechanics that it does not hurt to see it multiple times. The complete set of steps involved in solving any quantum mechanics problem (for particles moving in a 1D, time-independent potential) are;

1. Solve the time-independent Schrödinger equation,

$$\left(-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V(x)\right)\psi(x) = E\psi(x), \tag{2.24}$$

mathematically an eigen value problem, for the set of eigen functions of position $\psi_n(x)$, and the corresponding eigen values E_n, associated with $V(x)$,

2. Use the particle's wavefunction at time $t = 0$ to evaluate the constant expansion coefficients c_n that uniquely specify "■how much of each eigen function $\psi_n(x)$ is "needed to make up" $\Psi(x, 0)$,

$$c_n = \int_{-\infty}^{\infty} \psi_n^*(x)\Psi(x, 0)dx, \tag{2.25}$$

and thus,

$$\Psi(x, 0) = \sum_{n=1}^{\infty} c_n \psi_n(x), \tag{2.26}$$

The complete solution of the problem is given by explicitly expressing the full time-dependent wavefunction in terms of its value at $t = 0$ as,

$$\Psi(x, t) = \sum_{n=1}^{\infty} \left(\int_{-\infty}^{\infty} \psi_n^*(x')\Psi(x', 0)dx' \right) \psi_n(x)e^{-i\frac{E_n}{\hbar}t}. \tag{2.27}$$

2.5 Example 2: The Dynamical Properties of a Non-Stationary State in an Infinite Square Well Potential

It is finally time to do something interesting!

Solve for the full wavefunction, $\Psi(x, t)$ for a particle of mass m in an infinite square well potential of width a, assuming that at $t = 0$, $\Psi(x, 0) \propto (\psi_1(x) + \psi_2(x))$, and study some of its properties.

1. Start by normalizing the wavefunction at $t = 0$.

2. Find the full solution for all x, t.

3. Sketch what the wavefunction and its squared magnitude look like at $t = 0$.

4. Is the squared magnitude of the wavefunction periodic in time? If so, what is its period?

5. Sketch the wavefunction and its squared magnitude at $t = \frac{\pi}{6\omega_1}, \frac{\pi}{3\omega_1}, \frac{2\pi}{3\omega_1}$

6. Find the expectation value of the particle's position at $t = \frac{\pi}{6\omega_1}, \frac{\pi}{3\omega_1}, \frac{2\pi}{3\omega_1}$

7. Find the expectation value of the particle's momentum at $t = \frac{\pi}{6\omega_1}, \frac{\pi}{3\omega_1}, \frac{2\pi}{3\omega_1}$

8. Reflect on these results

Insert link or reference to the Infinite Square Well.jnt file here.

This reflection should help build intuition around how to interpret particle-like behaviour from looking at how the "centroid" of $|\Psi(x,t)|^2$ moves in time. The wave/particle starts with most of its being squished up next to one potential barrier, it "relaxes" towards the centre but its momentum carries it over to the other side. The gradient of the potential is infinite at the walls of the well, and this infinite force forces the wave/particle to reflect and repeat the process. If you thought of the wave as a water wave in a 1D pool, then it is the centre of mass dynamics that mimics a particle trajectory.

2.6   Example 3: The Harmonic Oscillator Potential

Another important and practically relevant potential profile is that associated with classical harmonic oscillators (mass on a spring), where $V(x) = \frac{1}{2}\omega_0^2 mx^2$, where ω_0 is the natural frequency of oscillation.

It is useful to review how this problem is usually approached classically. Given some initial position, x_0 and velocity, v_0, the full trajectory obtained from solving Newton's equation has the particle executing oscillatory harmonic motion between two extreme deviations, x_{max} from the minimum in the potential ($x = 0$ here). At the extrema, the particle velocity is zero, and the location of the extrema can be obtained from conservation of energy: $\frac{1}{2}\omega_0^2 mx_{max}^2 = \frac{1}{2}\omega_0^2 mx_0^2 + \frac{1}{2}mv_0^2$. Since the wavefunction that the quantum mechanical treatment of the problem generates connects most directly to the probability distribution for where the particle might be found if it is in a given quantum state, it is interesting to calculate the histogram for the probability of finding the classical particle between $-x_{max}$ and $+x_{max}$.

How would one go about calculating this classical histogram?

Imaging that the experiment would involve taking a photograph with a fixed exposure time at many randomly chosen times, enough to uniformly sample the motion of the oscillator with good signal to noise. The number of snapshots that captured the particle in any of the equally-wide spatial bins would represent the ordinate of the histogram, and the bin-centre position would represent the abscissa. The relevant calculation would then be the amount
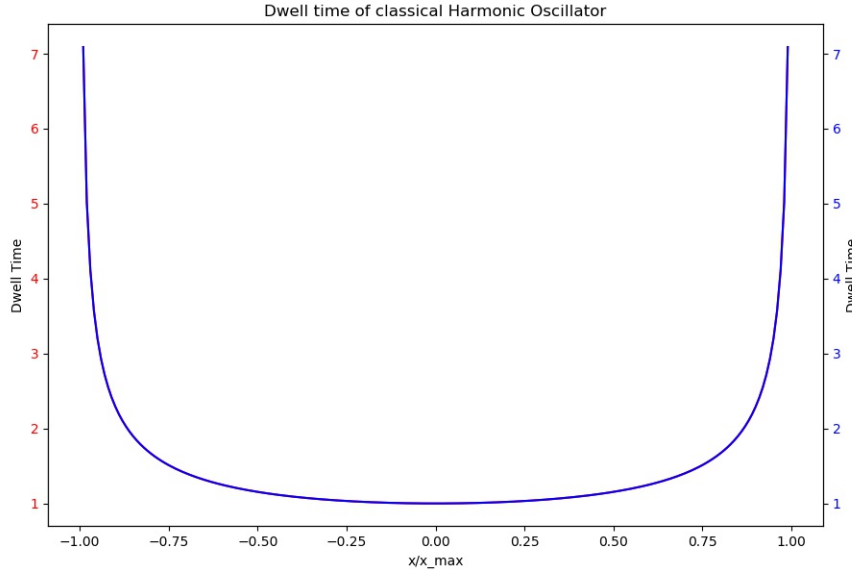
Figure 2.1: Calculated histogram of probability of finding a classical particle undergoing simple harmonic motion at any given location between its extrema (normalized to be unity at $x = 0$).

of time spent by the particle within a given bin's range (all the same), as a function of the central position of the bin. The time spent by the particle would be the bin width, δx, divided by the exposure time Δt, divided by the velocity, expressed in terms of position, $v(x)$, not time. The resulting function is

$$\frac{\Delta x}{\Delta t} \left| x_{max} \omega_0 sin(cos^{-1}(\frac{x}{x_{max}})) \right|^{-1} \tag{2.28}$$

and the resulting histogram is shown in Fig. 2.1.

A sanity check suggests this is qualitatively as one would expect, since the particle slows down at the extrema, and is going fastest at $x = 0$.

Now solve the problem using quantum mechanics. The first step is to write down the eigen value problem associated with the time-independent SE:

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}\omega_0^2 mx^2 \right) \psi(x) = E\psi(x) \tag{2.29}$$

This is obviously a much more difficult differential equation to solve than in the infinite square well problem, where $V(x)$ was zero. There are two distinctly different approaches to solving this problem. One is essentially a brute-force differential equation solving method that involves assuming a power series expansion of the eigen functions in the dimensionless parameter $\zeta = \sqrt{m\omega_0/\hbar}\, x$, which leads to a recursion relation for the coefficients of the

various powers of ζ for each distinct eigen function. This is covered in the textbook and in an Activity. Here, the use of ladder operators, a powerful accessory in the quantum mechanic's toolbox, is described. Obviously both approaches produce the same set of eigen functions!

First note that the time-independent SE can be written in this case as,

$$\frac{1}{2m} \left(\hat{p}^2 + (m\omega_0 x)^2 \right) \psi(x) = \hat{H}\psi(x) = E\psi(x), \tag{2.30}$$

where you should recall that we earlier found that $\hat{p} = -i\hbar\frac{d}{dx}$, a differential operator, and where $\hat{H}$ is introduced to represent the Hamiltonian, or total energy operator.

If this were the classical Hamiltonian, the transformation of variables

$$a_{\pm} \equiv \frac{1}{2\hbar m\omega_0}(\mp ip + m\omega_0 \ x) \tag{2.31}$$

would result in the very simple parameterization,

$$H = \hbar\omega_0 \ a_- a_+. \tag{2.32}$$

However, since $\hat{H}$ is a differential operator, it is **not** true that $\hat{H} = \hbar\omega_0 \ \hat{a}_- \hat{a}_+$. it Carefully expanding $\hat{a}_- \hat{a}_+$, one gets,

$$\hat{a}_- \hat{a}_+ = \hat{H} - \frac{i}{2\hbar} \left(x\hat{p} - \hat{p} \ x \right). \tag{2.33}$$

Obviously if $\hat{p}$ was not a differential operator, the last term would be zero and the classical result would be recovered. But the last term is not zero, and by operating on some arbitrary function $f(x)$ with the operator $(x\hat{p} - \hat{p} \ x)$, it can be shown that

$$(x\hat{p} - \hat{p} \ x) \equiv [x, \hat{p}] = i\hbar, \tag{2.34}$$

and therefore

$$\hat{H} = \hbar\omega_0 \left(\hat{a}_- \hat{a}_+ - \frac{1}{2} \right) = \hbar\omega_0 \left(\hat{a}_+ \hat{a}_- + \frac{1}{2} \right). \tag{2.35}$$

The operators $\hat{a}_+$ and $\hat{a}_-$ are referred to as raising and lowering ladder operators respectively, for reasons that will soon become clear.

If one assumes that there exists an eigen state of $\hat{H}$, call it $\psi_k(x)$ with some energy E_k, then the state obtained by operating on $\psi_k(x)$ with $\hat{a}_+$, turns out to also be an eigen state of $\hat{H}$, with an eigen value of $E_k + \hbar\omega_0$. Similarly, the state obtained by operating on $\psi_k(x)$ with $\hat{a}_-$ turns out to also be an eigen state of $\hat{H}$, with an eigen value of $E_k - \hbar\omega_0$. Neat trick! We will use a similar trick later when finding the eigen states of the hydrogen atom.

This then offers a method of generating all eigen states assuming you know one of them, and it also predicts exactly what their spacing is in energy. In order to pin down the exact energies, and the precise wavefunctions, it is necessary to recognize (see Problem 2.2 in the text) that if E was to be less than the minimum value of $V(x)$ for any x, (in this case zero),

then it would be impossible to normalize the wavefunction (it would necessarily diverge). Knowing that there must therefore be some lowest energy eigen state, label its eigen energy E_0, then it must be that $\hat{a}_-\psi_0(x) = 0 \times \psi_0(x)$, which implies that $\hat{H}\psi_0(x) = \frac{1}{2}\hbar\omega_0\psi_0(x)$, and therefore $E_0 = \frac{1}{2}\hbar\omega_0$, $E_1 = \frac{3}{2}\hbar\omega_0$, $\frac{5}{2}\hbar\omega_0$, etc. Almost finished! How would one go about finding the actual wavefunctions?

Using $\hat{a}_-\psi_0(x) = 0$ with the definition of $\hat{a}_-$ one gets the differential equation

$$\frac{d\psi_0(x)}{dx} = -\frac{m\omega_0}{\hbar}x\psi_0(x) \tag{2.36}$$

which is easily integrated to find

$$\psi_0(x) \propto e^{-\frac{m\omega_0}{2\hbar}x^2} \tag{2.37}$$

and the normalization factor can be easily deduced, so that

$$\psi_0(x) = \left(\frac{m\omega_0}{\pi\hbar}\right)^{\frac{1}{4}} e^{-\frac{m\omega_0}{2\hbar}x^2} \tag{2.38}$$

Done!

It should be clear how to generate all of the other eigen states from this ground eigen state using the differential form of the operator $\hat{a}_+$. Figure 2.2 shows several of the lowest order wavefunctions plotted atop the harmonic potential. Figure 2.3 shows a much higher energy eigen state associated with $n = 60$ What features of these wavefunctions are inconsistent with the classical picture of a particle with a energies E_n in a harmonic potential?

Note that there is some finite probability for the particle to be found beyond the "classical turning point". Also, there are many locations within the classically allowed region where there is zero probability of finding the quantum particle if it is in one of these stationary states. The particle is actually most likely to be found at $x = 0$ in the lowest order eigen state, while as n becomes large, if one averages over the fast oscillations, the probability distribution starts to resemble the classical histogram.

What features of these HO eigen states are similar to the infinite square well case, and how do they mostly differ?
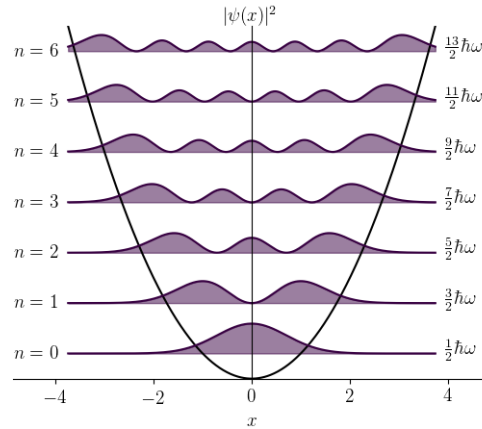
Figure 2.2: Plots of the lower order squared HO eigen functions superimposed on the harmonic potential (code from https://scipython.com/blog/the-harmonic-oscillator-wavefunctions/).
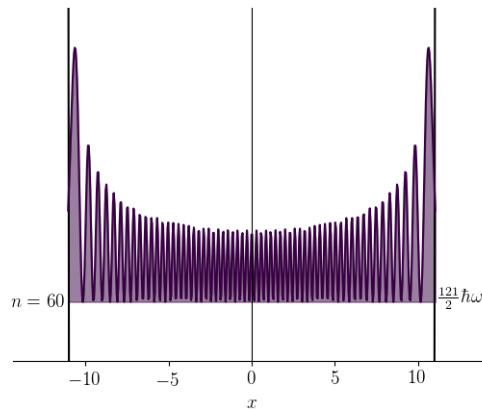

Figure 2.3: Plots of the n=60 squared HO eigen function (code from https://scipython.com/blog/the-harmonic-oscillator-wavefunctions/). Compare the average of this with the classical result in Fig. 2.1.

The number of nodes in the eigen states is equal to $n-1$, just as in the infinite well case. There are an infinite number of eigen states, also similar to the infinite well case, but now the spacing of the levels is equal, independent of n, whereas it increases with increasing n in the infinite well case.

2.7 Example 4: The Free Electron

The functional form of the solution of the time-indepdendent SE when $V(x) = 0$ has already been obtained when solving for the wavefunction inside the infinite square well potential. It is simply a harmonic plane wave function, or rather two harmonic plane waves associated with any energy eigen value E. There is a technical problem though, because there are no obvious boundary conditions that can be applied to localize the eigen states, and to therefore restrict the allowed eigen energies. It would appear that there is a continuum of energy eigen values that can be labelled by the magnitude of the wavevector of the harmonic wavefunction associated with that energy. It is tempting to therefore conclude that the two stationary states associated with an eigen energy $E(|k|)$ are,

$$\Psi_k(x,t) = e^{ikx}e^{-i\sqrt{\frac{2mE}{\hbar^2}}t}, \tag{2.39}$$

and

$$\Psi_{-k}(x,t) = e^{-ikx}e^{-i\sqrt{\frac{2mE}{\hbar^2}}t}. \tag{2.40}$$

It should be clear though, that there is a problem because these states cannot be normalized. Thus, technically, there are no valid stationary states of the free electron! Stated another way, there are no valid solutions of the free particle SE with a perfectly defined energy. This technicality is not of much practical importance when it is realized that it is still true that any valid wavefunction that satisfies the free particle SE can still be expressed as a superposition of all the separable mathematical solutions, but now the summation becomes an integral:

$$\Psi_{free}(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k)e^{i(kx-\frac{\hbar k^2}{2m}t)}dk \tag{2.41}$$

and importantly,

$$\Psi_{free}(x,0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k)e^{ikx}dk \tag{2.42}$$

The factor of $\frac{1}{2\pi}$ is a common convention associated with Fourier transform pairs: it should be recognized that $\phi(k)$ is precisely the Fourier component of the free particle wavefunction at $t = 0$. It must therefore be true that

$$\phi(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \Psi_{free}(x,0)e^{-ikx}dx. \tag{2.43}$$

Why does this offer a practical way out of having to deal with non-normalizable harmonic plane wave stationary states? It is easy to form a valid wavefuction that would look very

much like the nominal harmonic planewave free particle state $e^{i(k_0 x - \frac{\hbar k_0^2}{2m} t)}$ by using Eqn. 2.42 and choosing $\phi(k)$ to be a Gaussian function centred at $k = k_0$, with a very small variance, or small range of k values where $\phi(k)$ has any appreciable weight. Since the Fourier transform of a Gaussian is a Gaussian, with a variance inversely proportional to the original variance, then this state would effectively be the nominal harmonic plane wave, modulated by a Gaussian envelop function that would localize it (and thus solve the normalization problem), but if a sufficiently narrow range of k values is included in $\phi(k)$, that Gaussian envelope would extend over a very large distance in real space. Such a state is referred to as a wavepacket, and the point is that most of its important properties are determined only by k_0, and are not influenced by the exact extent of the localized envelope.

INSERT SOME VERSION OF Gaussian_wavepacket_start.jnt HERE

The free particle problem in quantum mechanics is therefore easy, but subtle. There are rigously no exact stationary states with precisely defined energy. The spectrum is continuous and exists only for $E > 0$. States with *almost* precisely defined energy $\frac{\hbar^2 k_0^2}{2m}$ and momentum $\hbar k_0$ consist of wavepackets with momentum components closely clustered around $k = k_0$, where the range of momentum values is dictated by the spatial extent of the exact wavefunction at $t = 0$. So long as that spatial extent is much much larger than $2\pi/k_0$, for all practical purposes the only impact of the precise spatial extent of the wavepacket is in determining the variance of the state's measured energy or momentum. The expectation value of momentum one would calculate for the pure harmonic planewave with wavevector k_0, by fudging the limits of integration, is $\hbar k_0$, and the expectation value of its energy using the same fudge would be $\frac{\hbar^2 k_0^2}{2m}$ as one might expect from the classical formula for the particle's kinetic energy. These expectation values are rigorously obtained using the wavepacket wavefunctions, where it is clear that the expectation value of momentum divided by the mass actually corresponds to the group velocity of the wavepacket, not the phase velocity, which is half the value of the group velocity. The group velocity is associated with the motion of the envelope of the wavepacket while the phase velocity is associated with the motion of the harmonic carrier.

2.8 Example 5: The Finite Square Well Potential

Using the potential $V(x)$ as sketched in Fig. 2.4, the algorithm for solving the problem is:

1. Solve for the general form of the eigen states of the time-independent SE in each of the three regions separately, and let the solution in each region be a linear combination of these allowed solutions with unknown coefficients.

2. Use the continuity of the wavefunction and its derivative at the boundaries to find the unknown coefficients, which will also dicatate the allowed eigen energies.

3. Recognize the fundamental differences in the nature of the eigen states obtained with eigen energies above and below the barrier potential.

It will become apparent that the nature of the solutions of the general equation are very different for energies $E > 0$ or $E < 0$. Starting with the case where $E < 0$, then it should be clear that the general form of the time-independent SE solution will be different in regions I, II, and III. If we use a label i to represent I, II, or III, then the general equation is

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V_i\right)\psi_i(x) = E\psi_i(x),\tag{2.44}$$

where $V_I = 0 = V_{III}$, and $V_{II} = -V_0$. In region I therefore,

$$\psi_I(x) = Ae^{-k_I x} + Be^{+k_I x}\tag{2.45}$$

with A, B constants, and with $k_I = +\frac{\sqrt{-2mE}}{\hbar}$. Since $E < 0$, the only way to have a normalizable solution is to choose $A = 0$, so that $\psi(x)_I = Be^{+\frac{\sqrt{-2mE}}{\hbar}x}$ remains finite as $x \to -\infty$. One immediately sees that in region III,

$$\psi_{III}(x) = Ee^{-k_{III}x} + Fe^{+k_{III}x}\tag{2.46}$$

and the same considerations imply $\psi_{III}(x) = Ee^{-\frac{\sqrt{-2mE}}{\hbar}x}$. In region II,

$$\psi_{II}(x) = Ce^{ik_{II}x} + De^{-ik_{II}x}\tag{2.47}$$

with $k_{II} = \frac{\sqrt{2m(E+V_0)}}{\hbar}$. Since region II is bounded, and the solutions harmonic, there is no normalizability restriction that would force C or D to be zero. There are therefore 4 unknown coefficients, B, C, D, E that must be determined by ensuring the continuity of the wavefunction and its derivative at the boundaries, which offers four equations: $\psi_I(-a) = \psi_{II}(-a), \psi_{II}(a) = \psi_{III}(a), \frac{d}{dx}\psi_I(-a) = \frac{d}{dx}\psi_{II}(-a)$, and $\frac{d}{dx}\psi_{II}(a) = \frac{d}{dx}\psi_{III}(a)$. The way in which satisfying these boundary conditions leads to an equation that restricts the allowed eigen values of E is insightful.

There are various mathematical ways to go about satisfying the four boundary conditions. One approach invokes a symmetry argument that takes advantage of the fact that the potential $V(x)$ is mirror symmetric about $x = 0$, so that $V(-x) = V(x)$ This implies that if $\psi(x)$ is a solution of the time-independent SE for this potential, then so is $\psi(-x)$. This in turn implies that each solution must either be symmetric $\psi(-x) = \psi(x)$ or anti-symmetric $\psi(-x) = -\psi(x)$ (think about it!). Using this property the symmetric solutions can be obtained by noting that for these, $B = E, C = D$, while for the antisymmetric solutions, $E = -B, D = -C$. Then there are only 2 equations that must be solved for each of the symmetric and antisymmetric cases separately.

In the symmetric case the boudary conditions at $x = -a$ reduce to,

$$Be^{-k_I a} = 2C\cos(k_{II}a),\, Bk_I e^{-k_I a} = 2k_{II}C\sin(k_{II}a).\tag{2.48}$$

Dividing these two equations leads to

$$tan(k_{II}a) = \frac{k_I}{k_{II}}.\tag{2.49}$$

Substituting for k_I and k_{II} in terms of E and V_0 then indicates that all even solutions of the time-independent SE for the finite potential well with $E < 0$ must have eigen energies that satisfy the transcendental equation

$$tan\left(\sqrt{\frac{2m(E+V_0)}{\hbar^2}}a\right) = \sqrt{\frac{-E}{E+V_0}}. \tag{2.50}$$

In the anti-symmetric case the boudary conditions at $x = -a$ reduce to,

$$Be^{-k_I a} = -2iC\sin(k_{II}a), \, Bk_I e^{-k_I a} = 2iCk_{II}C\cos(k_{II}a). \tag{2.51}$$

Dividing these two equations leads to

$$\tan(k_{II}a) = -\frac{k_{II}}{k_I}. \tag{2.52}$$

Substituting for k_I and k_{II} in terms of E and V_0 then indicates that all odd solutions of the time-independent SE for the finite potential well with $E < 0$ must have eigen energies that satisfy the transcendental equation

$$\tan\left(\sqrt{\frac{2m(E+V_0)}{\hbar^2}}a\right) = -\sqrt{\frac{E+V_0}{-E}}. \tag{2.53}$$

Such transcendental equations must be solved numerically, or graphically (see text for the graphical approach). Once each possible eigen value is obtained, the corresponding eigen functions can be obtained by solving for B/C from the appropriate boundary condition equations, and then normalizing the resulting function.

Interesting properties of these solutions for $E < 0$ include;

1. as should be intuitively clear from our solution of the infinite potential well case, the number of solutions depends on the depth of the potential well (more solutions the deeper the well)

2. not so obvious is the fact that there is always at least one allowed eigen value (necessarily symmetric)

Now look at the solutions when $E > 0$. In regions I and III, if one defines $k_I = k_{III} = \sqrt{\frac{2mE}{\hbar^2}}$, then

$$\psi_I(x) = Ae^{ik_I x} + Be^{-ik_I x} \tag{2.54}$$
$$\psi_{III}(x) = Ee^{ik_{III}x} + Fe^{-ik_{III}x} \tag{2.55}$$

and still

$$\psi_{II}(x) = Ce^{ik_{II}x} + De^{-ik_{II}x}, \tag{2.56}$$

with k_I, k_{II}, k_{III} all real, so the solutions in all regions are now harmonic plane waves (free-electron like). Recall that there was a technical problem associated with the fact that these

harmonic functions do not tend to zero at $\pm\infty$, and so are not formally normalizable. But as in the free electron case we can proceed by ignoring this and thinking in terms of wavepackets made up of a collection of solutions about a narrow range of k values.

To appreciate the qualitative difference in the solutions for a free electron, and for the electrons with $E > 0$ that encounter a finite potential well, first solve the simpler problem of an electron encountering a potential step (half an infinite well, as shown in Fig. 2.5).

Following the solution algorithm above means it is necessary only to find the A, B, C, D complex coefficients that satisfy $\psi_I(-a) = \psi_{II}(-a)$ and $\frac{d}{dx}\psi_I(-a) = \frac{d}{dx}\psi_{II}(-a)$. But there are only two equations and four unknowns! To understand how to resolve this apparent problem, recall that in the free electron case there were actually two distinct, orthogonal eigen state solutions of the time-independent SE for any $E > 0$, $e^{i\sqrt{\frac{2mE}{\hbar^2}}x}$ and $e^{-i\sqrt{\frac{2mE}{\hbar^2}}x}$, corresponding to full solutions of the SE of $e^{i(\sqrt{\frac{2mE}{\hbar^2}}x - \frac{E}{\hbar}t)}$ and $e^{i(-\sqrt{\frac{2mE}{\hbar^2}}x - \frac{E}{\hbar}t)}$. These respectively correspond to rightward and leftward propagating harmonic plane wavefunctions. They are not normalized or normalizable, but one doesn't worry about that because eventually they would become normalized wavepackets that behave almost identically to these solutions. It is therefore reasonable to anticipate that there will be two distinct solutions of this potential step problem for every $E > 0$, one corresponding to a rightward propagating solution, and the other to a leftward propagating solution.

Physically one can associate this to the classical electrodynamic problem of a light beam incident from say the left, on a very thick piece of glass. What happens there is that some of the light reflects, and some is transmitted, and the properties of the solution are completely captured by evaluating the reflection and transmission coefficients of the electric field associated with the light field. In our potential step problem, to find this rightward propagating solution, one sets $D = 0, A = 1, B = R, C = T$, where the normalization is ignored (just like for free electrons and the optical problem), and B therefore becomes the reflectivity R, and E becomes the transmission, T. This solves our problem, since now there are only two unknowns and we have two equations.

If there were a light source embedded within the glass, illuminating the interface from the right, then there would be a leftward propagating wave exiting into the air, and a reflected beam going to the right inside the glass. The analogous leftward propagating solutions of the potential step problem are then obtained by satisfying the boundary conditions assuming $A = 0, B = T', C = R', D = 1$, where in general the reflection and transmission coefficients at a given energy eigen value $E > 0$ will be different for leftward and rightward solutions.

For the rightward propagating case, the boundary conditions at $x = -a$ lead to the equations,

$$e^{-ik_I a} + Re^{ik_I a} = Te^{-ik_{II} a} \tag{2.57}$$

$$ik_I\left(e^{-ik_I a} - Re^{ik_I a}\right) = ik_{II}Te^{-ik_{II} a}, \tag{2.58}$$

from which one immediately finds

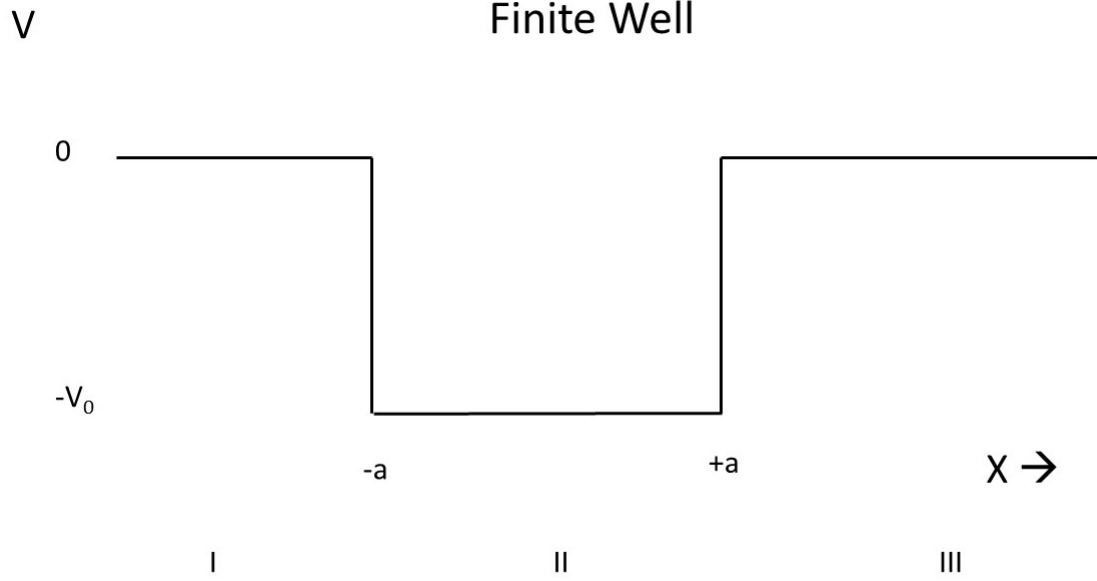$$R = \frac{k_I - k_{II}}{k_I + k_{II}}e^{-2ik_I a} \tag{2.59}$$

Figure 2.4: The finite potential well.

$$T = \frac{2k_I}{k_I + k_{II}} e^{(k_{II} - k_I)a}, \tag{2.60}$$

which implies

$$\psi_{right}(x) = \left(e^{i\sqrt{\frac{2mE}{\hbar^2}}x} + \left(\frac{\frac{\sqrt{E}}{E+V_0} - 1}{\frac{\sqrt{E}}{E+V_0} + 1} \right) e^{-i\sqrt{\frac{2mE}{\hbar^2}}x} \right)(1 - \Theta(x+a)) \tag{2.61}$$

$$+ \left(\frac{2\sqrt{E}}{\sqrt{E} + \sqrt{E+V_0}} \right) e^{i\sqrt{\frac{2m(E+V_0)}{\hbar^2}}x} \Theta(x+a) \tag{2.62}$$

where $\Theta(x)$ is the Heavyside function.

Note that imposing the boundary conditions in this case does not restrict E so long as it is positive. Following the same approach, the leftward propagating solution at energy E is

$$\psi_{left}(x) = \left(e^{-i\sqrt{\frac{2m(E+V_0)}{\hbar^2}}x} + \left(\frac{1 - \frac{\sqrt{E}}{E+V_0}}{1 + \frac{\sqrt{E}}{E+V_0}} \right) e^{2i\sqrt{\frac{2m(E+V_0)}{\hbar^2}}a} e^{i\sqrt{\frac{2m(E+V_0)}{\hbar^2}}x} \right)\Theta(x+a) \tag{2.63}$$

$$+ 2\frac{2\sqrt{E+V_0}}{\sqrt{E} + \sqrt{E+V_0}} e^{i\left(\sqrt{\frac{2m(E+V_0)}{\hbar^2}} - \sqrt{\frac{2mE}{\hbar^2}}\right)a} e^{-i\sqrt{\frac{2m(E)}{\hbar^2}}x}(1 - \Theta(x+a)) \tag{2.64}$$

Returning to the full finite potential well potential problem, the analogous approach is then to solve for the rightward eigen states by setting $A = 1, B = R, E = T$, and solving
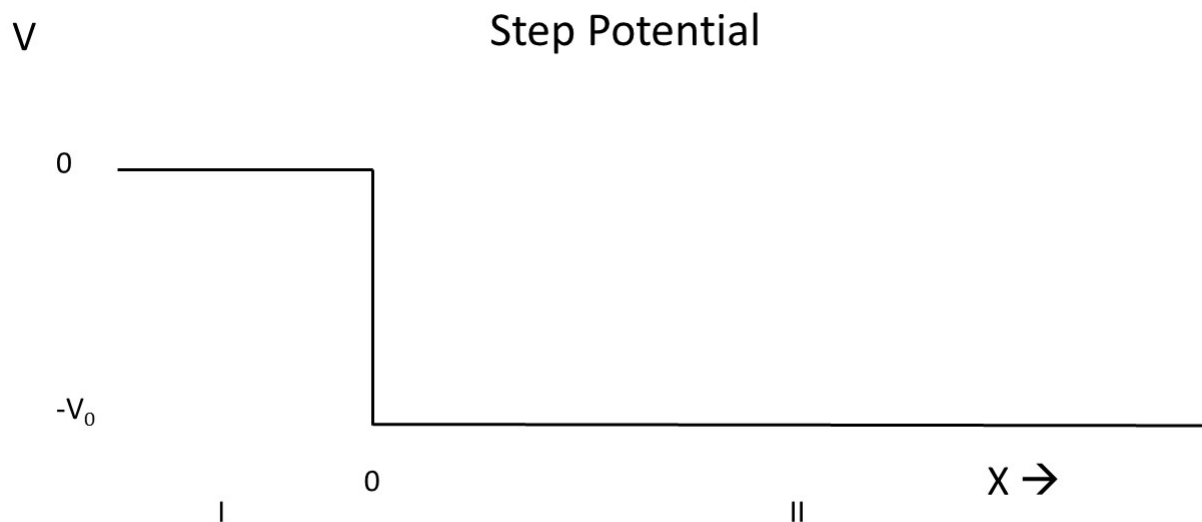
Figure 2.5: The single step potential profile.

for R, T, C, D using the four boundary conditions at $x = \pm a$. Note that one can no longer use the reflection symmetry of the potential about $x = 0$ to simplify the solution since the symmetry is broken by assuming a rightward-propagating solution. It is still useful though, since once you have the rightward propagating solution, you automatically know the leftward propagating solution. The solutions for R and T are given in the textbook. Again, there is no restriction on the value of E so long as it is positive.

Chapter 3

More Formal, More General Formulations of Quantum Mechanics

3.1 Abstract State Space

The introductory chapters were designed to provide a heuristic transition from classical to quantum mechanics. Since classical mechanics (of particle motion) is all based on trajectories $x(t)$, it was natural to emphasize the role of the wavefunction $\Psi(x,t)$ in quantum mechanics, as it arguably has the most direct connection to trajectories, through the interpretation of $|\Psi(x,t)|^2$ as representing a probability distribution for where to "find the particle" at any given time. However, once the basic concepts, that make quantum mechanics so radically different than classical mechanics, are grasped, and one gains experience solving real quantum mechanical problems, it quickly becomes apparent that the utility of the particle's wavefunction is actually quite limited in comparison to the much more general "state space" interpretation of the quantum state that carries all of the information possible about a particle's dynamics.

This chapter covers the most challenging material in the course. This is because the "state space" view of quantum mechanics is totally abstract: it will become apparent that the properties of the particle's immutable, abstract *state* can only be "interpreted" by "projecting" that state onto one of an arbitrarily large number of "basis" states that do have some physical connection. Again, as will become apparent in this chapter, the wavefunction $\Psi(x,t)$ actually specifies how much of the system's immutable state projects onto the "position basis" at time t. For the purpose of foreshadowing, there is absolutely no more information carried by $\Psi(x,t)$ than there is by a function $\Phi(p,t)$ that will be seen to represent how much of the same immutable state projects onto the "momentum basis" at time t. And those are only two of a multitude of possible bases onto which the *state* can be projected.

From the preceeding it should be clear that the terminology surrounding state space and its interpretation bears a strong similarity to that used in linear algebra. This is no accident, and for that reason the next section will review concepts that should be familiar from linear algebra in a 2D real vector space. This will serve as a useful vehicle for introducing the Dirac

"bra ket" notation used to manipulate a particle's state in abstract state space.

3.2 Review of Linear Algebra in 2D Real Vector Space

Two dimensional real vectors are readily rendered using easily interpreted graphs, and their linear algebra is almost sufficent to immediately generalize to the abstract Hilbert space of a quantum particle. Recall that on a flat 2D map, if one is in Vancouver and one wants to precisely specify the shortest path to the North Pole, it would simply be a straight line connecting two dots on the map. If one put an arrow on the end located at the north pole, it would be a 2D real vector that defined the distance (length of the vector) and the direction (due north) one would take to travel from Vancouver to the North Pole. The same line with an arrow at the Vancouver end would represent the distance and direction one would travel to get from the North Pole to Vancouver (the length is the same but the direction is opposite, so they are distinct vectors).

Now consider how to convey the properties of this immutable vector to the traveller in Vancouver. Clearly it would be natural to "define it" by specifying that it was made up of a distance x oriented towards the east, and a distance y oriented towards the north, from that point. This all is clearly represented graphically as in Fig. **??** where the vector connecting Vancouver to the North Pole is labelled $\vec{r}$, and it is the addition of a vector $\vec{r}_E$ and $\vec{r}_N$. The length of $\vec{r}_E$ is "the projection of " the vector $\vec{r}$ onto a unit length "basis vector" $\hat{i}$ that always points eastward, and the length of $\vec{r}_N$ is the "projection of " the vector $\vec{r}$ onto a unit length "basis vector" $\hat{j}$ that always points northward. The unit vectors are used to specify direction, and they can be multiplied by real scaler quantitites (the projections) that convey the length. Using linear albegraic notation,

$$\vec{r} = r_E \hat{i} + r_N \hat{j}. \tag{3.1}$$

But note that these unit vectors are chosen by convention, a convenient convention! Mathematically there is no reason that two unit vectors $\hat{\alpha}_1 = \sqrt{\frac{1}{2}}\left(\hat{i} + \hat{j}\right)$, and $\hat{\alpha}_2 = \sqrt{\frac{1}{2}}\left(-\hat{i} + \hat{j}\right)$ pointing respectively to the north east and north west could not be chosen to specify the unchanging, immutable vector $\vec{r}$. Thus

$$\vec{r} = r_E \hat{i} + r_N \hat{j} = r_{NE}\hat{\alpha}_1 + r_{NW}\hat{\alpha}_2. \tag{3.2}$$

Clearly there are an infinite number of distinct orthogonal unit (basis) vectors that could be used to "project $\vec{r}$ onto" in order to unambiguously specify it, but for any of those distinct bases, the "projection of $\vec{r}$" onto each of the two basis vectors is unique:

$$r_E = \vec{r} \cdot \hat{i} = |\vec{r}|cos(\phi), etc. \tag{3.3}$$

where ϕ is the angle subtended by $\vec{r}$ and $\hat{i}$. This projection is referred to as the "dot product", or more generally, the "inner product" of the two vectors.

Given a defined set of ordered orthogonal unit basis vectors, then any arbitrary vector $\vec{r}$ can be algebraically defined in terms of its projections along the first and second unit vectors as a tuple, for instance

$$\vec{r} \equiv \begin{bmatrix} r_E \\ r_N \end{bmatrix} \tag{3.4}$$

in the ordered $\hat{i}, \hat{j}$ basis. *As an aside, it would actually make sense to use the following notation, to make the implied basis explicit:*

$$\vec{r} \equiv \begin{bmatrix} r_E \\ r_N \end{bmatrix}_{\hat{i},\hat{j}}, \tag{3.5}$$

but this is not typically used. The dot or inner product of two vectors $\vec{r}$ and $\vec{r}\,'$

$$\vec{r} \cdot \vec{r}\,' = r_E r'_E + r_N r'_N = |\vec{r}||\vec{r}\,'|cos(\phi) \tag{3.6}$$

is denoted in column vector notation as

$$\vec{r} \cdot \vec{r}\,' = \begin{bmatrix} r_E & r_N \end{bmatrix} \begin{bmatrix} r'_E \\ r'_N \end{bmatrix} \tag{3.7}$$

where

$$\begin{bmatrix} r_E & r_N \end{bmatrix} \tag{3.8}$$

is referred to as the transpose of the vector

$$\begin{bmatrix} r_E \\ r_N \end{bmatrix}, \tag{3.9}$$

or explicitly,

$$\begin{bmatrix} r_E & r_N \end{bmatrix} = \begin{bmatrix} r_E \\ r_N \end{bmatrix}^T. \tag{3.10}$$

It is now time to introduce Dirac's bra ket notation. A ket corresponding to a 2D real vector is simply denoted

$$|\vec{r}\rangle \equiv \begin{bmatrix} r_E \\ r_N \end{bmatrix}. \tag{3.11}$$

To motivate this notation, notice that the inner product of $\vec{r}$ and $\vec{r}\,'$ is now

$$\vec{r} \cdot \vec{r}\,' = \langle \vec{r}||\vec{r}\,'\rangle \equiv \langle \vec{r}\,|\vec{r}\,'\rangle, \tag{3.12}$$

and more useful, one can express the decomposition of an arbitrary vector in a particular orthogonal basis as

$$|\vec{r}\rangle = \langle \hat{i}\,|\hat{r}\rangle|\hat{i}\rangle + \langle \hat{j}\,|\hat{r}\rangle|\hat{j}\rangle \tag{3.13}$$

and since $\langle \hat{i}\,|\hat{r}\rangle$ and $\langle \hat{j}\,|\hat{r}\rangle$ are just numbers, this can be written

$$|\vec{r}\rangle = |\hat{i}\rangle\langle \hat{i}\,|\hat{r}\rangle + |\hat{j}\rangle\langle \hat{j}\,|\hat{r}\rangle = \left(|\hat{i}\rangle\langle \hat{i}| + |\hat{j}\rangle\langle \hat{j}| \right) |\vec{r}\rangle. \tag{3.14}$$

Clearly the object in parentheses must be just the number 1? No!! It is the unit 2 x 2 matrix, which is why this notation is so useful. If one goes back to the cuple notation,

$$|\hat{i}\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, |\hat{j}\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{3.15}$$

then

$$|\hat{i}\rangle\langle\hat{i}| \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix}. \tag{3.16}$$

How does one interpret this? If one returns to Eqn. 3.14, the answer is clear, since the term $\langle\hat{i}\,|\hat{r}\rangle|\hat{i}\rangle$ is meant to render the (vector) component of $\vec{r}$ in the $\hat{i}$ direction, which is clearly a vector. Rearranging the order of the number $\langle\hat{i}\,|\hat{r}\rangle$ and the vector $|\hat{i}\rangle$ can't change the fact that the result is a vector, thus

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} \equiv \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{3.17}$$

and therefore,

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{3.18}$$

This is referred to as dyadic or direct project notation in linear algebra. The rule for converting

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \end{bmatrix}. \tag{3.19}$$

into a matrix is

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \end{bmatrix} = \begin{bmatrix} a_1 b_1 & a_1 b_2 \\ a_2 b_1 & a_2 b_2 \end{bmatrix} \equiv |\vec{a}\rangle\langle\vec{b}|. \tag{3.20}$$

It should be clear now why the quantity $|\hat{i}\rangle\langle\hat{i}| + |\hat{j}\rangle\langle\hat{j}|$ is referred to as *the projection operator in the $\hat{i}, \hat{j}$ basis*, since when it acts on some arbitrary vector (or ket), each component of the operator acting on the vector renders the component of that vector in the chosen basis, and together all the components are summed to yield the original vector. Note that we could use any orthonormal basis, so $|\hat{\alpha}_1\rangle\langle\hat{\alpha}_1| + |\hat{\alpha}_2\rangle\langle\hat{\alpha}_2|$ is the projection operator in the $\hat{\alpha}_1, \hat{\alpha}_2$ basis.

This notion of an *operator* that *operates on* vectors in 2D real vector space should be familiar from introductory linear algebra, where the *rotation operator* would have been introduced. Recall there that if the coordinates of some vector $\vec{r}$ (or in bra ket notation $|\vec{r}\rangle$), are x, y in the $\hat{i}, \hat{j}$ basis ($\begin{bmatrix} x \\ y \end{bmatrix}$ in vector notation), then the *matrix operator*, $\overline{R}(\theta)$ that, when it multiplies $\begin{bmatrix} x \\ y \end{bmatrix}$, yields the coordinates of the vector $\vec{r}_\theta$ (again in the $\hat{i}, \hat{j}$ basis) that is rotated by an angle θ with respect to $\vec{r}$. Then

$$\begin{bmatrix} x_\theta \\ y_\theta \end{bmatrix} = \overline{R}(\theta) \begin{bmatrix} x \\ y \end{bmatrix} \tag{3.21}$$

or in bra ket notation,

$$|\vec{r}_\theta\rangle = \hat{R}(\theta)|\vec{r}\rangle \tag{3.22}$$

where $\hat{R}(\theta)$ is the *rotate by angle θ operator in the $\hat{i}, \hat{j}$ basis,* and $\overline{R}(\theta)$ is its matrix representation. It is fairly easy to show that in 2D real vector space, the matrix representation of the *rotate by angle θ* operator is

$$\hat{R}(\theta) \equiv \begin{bmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{bmatrix} \tag{3.23}$$

and it is independent of the basis used to represent or expand the vectors it operates on. More generally this is not the case, and the matrix representation of an operator must in general be allowed to depend on the basis. In bra ket notation this very important concept can be understood as follows,

$$|\vec{r}_\theta\rangle = \hat{R}(\theta)|\vec{r}\rangle = \hat{R}(\theta)\left(\sum_{j=1}^{2} |\hat{\alpha}_j\rangle\langle\hat{\alpha}_j|\right)|\vec{r}\rangle \tag{3.24}$$

where the unity operator has been inserted between the physical operator of interest, and the ket it acts upon. **This "insertion of the unity operator in any chosen basis" is a key concept that you MUST thoroughly understand**. By projecting the basis state $|\hat{\alpha}_i\rangle$ onto the above equation, one gets an explict expression for the $\hat{alpha}_i^{th}$ projection of the final vector/ket, that results from the operator operating on the components of the initial vector/ket, also expanded in the same $\hat{\alpha}_i, \hat{\alpha}_j$ basis used to expand the unity operator in the general expression:

$$\langle\hat{\alpha}_i|\vec{r}_\theta\rangle = \sum_{j=1}^{2}\langle\hat{\alpha}_i|\hat{R}(\theta)|\hat{\alpha}_j\rangle\left(\langle\hat{\alpha}_j|\vec{r}_\theta\rangle\right) \tag{3.25}$$

To make the connection to matrix representations, this is equivalent to

$$\begin{bmatrix} x_\theta \\ y_\theta \end{bmatrix}_{\hat{\alpha}_i,\hat{\alpha}_j} = \overline{R}(\theta)_{\hat{\alpha}_i,\hat{\alpha}_j} \begin{bmatrix} x \\ y \end{bmatrix}_{\hat{\alpha}_i,\hat{\alpha}_j}. \tag{3.26}$$

To see how $\langle\hat{\alpha}_i|\hat{R}(\theta)|\hat{\alpha}_j\rangle \equiv \hat{R}(\theta)_{i,j}$, *the matrix elements of the rotation operator* are interpreted, consider $\langle\hat{\alpha}_i|\hat{R}(\theta)|\hat{\alpha}_j\rangle$ re-written as $\langle\hat{\alpha}_i|\left(\hat{R}(\theta)|\hat{\alpha}_j\rangle\right)$. It should be clear that it therefore corresponds to "rotate basis vector $\hat{\alpha}_j$ by angle θ, and then project basis vector $\hat{\alpha}_i$ onto the result". With reference to the vector diagram in Fig. 3.1 it should be clear that for any arbitrary $\hat{\alpha}_i, \hat{\alpha}_j$ basis,

$$\begin{aligned}
\langle\hat{\alpha}_1|\hat{R}(\theta)|\hat{\alpha}_1\rangle &= cos(\theta) & (3.27) \\
\langle\hat{\alpha}_2|\hat{R}(\theta)|\hat{\alpha}_1\rangle &= sin(\theta) & (3.28) \\
\langle\hat{\alpha}_1|\hat{R}(\theta)|\hat{\alpha}_2\rangle &= -sin(\theta) & (3.29) \\
\langle\hat{\alpha}_2|\hat{R}(\theta)|\hat{\alpha}_2\rangle &= cos(\theta), & (3.30)
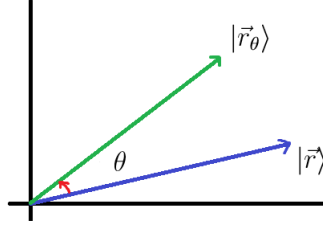\end{aligned}$$

Figure 3.1: Rotation of a vector by

$$\hat{R}(\theta)$$

.

or,

$$\hat{R}(\theta)_{\hat{\alpha}_1,\hat{\alpha}_2} = \begin{bmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{bmatrix},$$ (3.31)

which proves the claim above that the matrix representation of the rotation operator in 2D real vector space is independent of the chosen basis.

As a next step in getting familiar with bra ket notation, and the concept of operators, consider the bra ket notation for determining the projection of an immutable vector $\vec{r}\,'$ onto an immutable vector obtained by rotating an immutable vector $\vec{r}$ by an angle θ, assuming that both the vectors and the matrix representation of the rotation operator are to be evaluated in the $\hat{\alpha}_i, \hat{\alpha}_j$ basis:

$$\langle \vec{r}\,'|\hat{R}(\theta)|\vec{r}\rangle = \sum_{i=1}^{2}\sum_{j=1}^{2}\langle \vec{r}\,'|\hat{\alpha}_i\rangle\langle\hat{\alpha}_i|\hat{R}(\theta)|\hat{\alpha}_j\rangle\langle\hat{\alpha}_j|\vec{r}\,'\rangle.$$ (3.32)

Note that this was obtained by simply inserting the unity operator expressed in the $\hat{\alpha}_i, \hat{\alpha}_j$ basis on either side of the operator. **Again, make sure you thoroughly understand this concept!**. Again recall that you can think of the right hand side of this equation in matrix form, as:

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\langle \vec{r}\,'|\hat{\alpha}_i\rangle\langle\hat{\alpha}_i|\hat{R}(\theta)|\hat{\alpha}_j\rangle\langle\hat{\alpha}_j|\vec{r}\,'\rangle = \begin{bmatrix} x\,' & y\,' \end{bmatrix}_{\hat{\alpha}_i,\hat{\alpha}_j} \overline{R}(\theta)_{\hat{\alpha}_i,\hat{\alpha}_j} \begin{bmatrix} x \\ y \end{bmatrix}_{\hat{\alpha}_i,\hat{\alpha}_j}.$$ (3.33)

For the same immutable vectors $\vec{r}$ and $\vec{r}\,'$, one could ask what bra ket equation represents projecting the vector obtained by rotating $\vec{r}\,'$ by an angle ϕ, onto the vector $\vec{r}$? The rotation operator on the vector $\vec{r}\,'$ by angle ϕ in bra ket notation is

$$|\vec{r}\,'_\phi\rangle = \hat{R}(\phi)|\vec{r}\,'\rangle \equiv \begin{bmatrix} x\,'_\phi \\ y\,'_\phi \end{bmatrix}_{\hat{\alpha}_i,\hat{\alpha}_j}$$ (3.34)

in some basis $\hat{\alpha}_i, \hat{\alpha}_j$. To project this rotated vector onto $\vec{r}$, the ket $|\vec{r}\,'_\phi\rangle$ must be converted into the bra $\langle \vec{r}\,'_\phi| = \langle \left(\hat{R}(\phi)|\vec{r}\,'\rangle \right)|$. To understand how to do this conversion, it is again useful to revert to the matrix representation where for any general real matrix $\overline{M}$, and real vector $\vec{v}$, the transpose of the vector $\overline{M}\vec{v}$ (equivalent to the bra version of the ket $\hat{M}|\vec{v}\rangle$)) is clearly $\left(\overline{M}\vec{v} \right)^T$, where $()^T$ represents the transpose operation. it is generally true that

$$\left(\overline{M}\vec{v} \right)^T = \vec{v}^T \overline{M}^T, \tag{3.35}$$

where obviously the order of the operations is critical. From this it should be clear that

$$\langle \vec{r}\,'_\phi| = \langle \vec{r}\,'|\hat{R}(\phi)^T. \tag{3.36}$$

In order for the projection of this rotated $\vec{r}\,'$ onto $\vec{r}$ to be the same as that obtained by first rotating $\vec{r}$ by angle θ, and then projecting $\vec{r}\,'$ onto it, then

$$\left(\langle \vec{r}\,'|\hat{R}(\phi)^T \right) |\vec{r}\rangle = \langle \vec{r}\,'|\hat{R}(\phi)^T|\vec{r}\rangle = \langle \vec{r}\,'|\hat{R}(\theta)|\vec{r}\rangle, \tag{3.37}$$

so $\hat{R}(\phi)^T = \hat{R}(\theta)$, which means that ϕ must equal $-\theta$, since, independent of the chosen basis, this implies, in the matrix representation, that

$$\hat{R}(\phi)^T \equiv \begin{bmatrix} cos(\phi) & sin(\phi) \\ -sin(\phi) & cos(\phi) \end{bmatrix} = \hat{R}(\theta) \equiv \begin{bmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{bmatrix}. \tag{3.38}$$

This exercise may have seemed somewhat abstract, but it turns out to be a very important concept/tool in the quantum mechanic's toolbox. The key takeaway message is that

If an operator, $\hat{O}$, (which can always be defined in terms of some "matrix" with respect to some basis), acts on an arbitrary vector, $|\vec{r}\rangle$, to produce a new vector $\hat{O}|\vec{r}\rangle$, that yields a certain value when some other arbitrary vector $|\vec{r}\,'\rangle$ is projected onto it, then the same value will be obtained if the vector generated by operating on $|\vec{r}\,'\rangle$ by the transpose of the original operator, $\hat{O}^T$ is projected onto the original $|\vec{r}\rangle$. This is succinctly summarized in bra ket notation by the important equation:

$$\langle \vec{r}\,'| \left(\hat{O}|\vec{r}\rangle \right) = \langle \left(\hat{O}^T|\vec{r}\,'\rangle \right) |\vec{r}\rangle = \langle \vec{r}\,'|\hat{O}|\vec{r}\rangle, \tag{3.39}$$

that applies for a general, physically realizable operator $\hat{O}$ acting in a real vector space.

The power of this bra ket notation, even when applied in this relatively simple 2D vector space example, lies in the fact that it is much more general than the arguably simpler mathematical matrix formula:

$$\begin{bmatrix} x' & y' \end{bmatrix} \left(\begin{bmatrix} O_{1,1} & O_{1,2} \\ O_{2,1} & O_{2,2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right) = \left(\begin{bmatrix} O_{1,1} & O_{1,2} \\ O_{2,1} & O_{2,2} \end{bmatrix}^T \begin{bmatrix} x' \\ y' \end{bmatrix} \right)^T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x' & y' \end{bmatrix} \begin{bmatrix} O_{1,1} & O_{1,2} \\ O_{2,1} & O_{2,2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$
$$\tag{3.40}$$

The bra ket version can be generally interpreted as a defining equation for an operator $\hat{O}^T$, call it the *dual operator* of the operator $\hat{O}$, without any reference to matrix representations at all. The pure graphical analysis done in the previous section could then be used to conclude that the *dual operator* of the "rotate by angle θ" operator, is "rotate by angle $-\theta$", without using any mathematical property of the matrix representation of these vectors and the physical operations that can be performed on them. The *less general* result is then that when a specific coordinate system is adopted, and column vectors and matrices are used to represent the physical vectors and operations, then it turns out that the matrix representation of the *dual operator* of any physically realizable operator is the transpose of the matrix representation of the operator itself.

3.3 The Connection to Quantum Mechamics

A summary of the most relevant elements from the linear algebra in 2D real vector spaces includes:

1. there exist immutable vectors that represent useful physical quantitites

2. to unambiguously specify any such vector, it is necessary to define a coordinate system (not immutable, there are in general an infinite number of possible coordinate systems) that consists of a set of orthogonal vectors of unit length, one for each dimension of the space

3. if the set of unit vectors is complete, then any arbitrary immutable vector can be uniquely defined by specifying the 'amount of' each unit vector that is required to include in a sum of weighted unit vectors that equals the original vector

4. the 'amount of' a given unit basis vector contained in any immutable vector is determined by the dot or inner product of the immutable vector and the unit vector(s)

5. if the set of orthogonal unit vectors is put into a well-defined order, then any immutable vector is readily specified by an ordered set of the dot-products along each of the unit vectors. if arranged in a vertical column, this ordered set of dot products represents the vector's coordinates in that coordinate system

6. there is typically a 'natural' set of unit vectors (eg. East and North) that is used to do real calculations: any of the other possible coordinate systems and basis vectors can always be expressed in terms of this natural set of unit vectors

7. in 2D real vector space it may not be obvious why one would use any other coordinate system, but this flexibility is very important in quantum mechanics

8. various 'operators' can be defined that, when applied to some vector, modify that vector in some way. The rotation operator is a simple, intuitive example in 2D real vector space

9. operators can be represented by matrices, and the result of the 'operation' on a given vector is obtained by multiplying its column coordinate vector from the left by the 'operator matrix'

10. the elements of each row of the matrix representing a given operator are obtained by left-multiplying each of the 'input' unit vectors of the assumed coordinate system (that specify the row number), one by one, and taking the dot product of the resulting 'operated on' unit vector by left-multiplying it by each of the transposed 'output' unit vectors in the basis (that then specify the column in that row of the matrix)

11. the matrix representing a given operator can also be obtained by first left-multiplying each of the 'output' unit vectors by the *transpose* of the matrix representing the operation, taking the transpose of each of the resulting vectors, and using them to dot product with each of the 'input' unit column vectors

12. while the components of a given immutable vector will necessarily change if the coordinate system changes, the matrix elements representing a given physical operator may or may not depend on the coordinate system used

To motive the connection of this to quantum mechanics, recall that the expectation value of a particle's momentum and position could be cast in the form

$$m < v > (t) \;\; = \;\; \int_{-\infty}^{+\infty} \Psi^*(x,t) \left(-i\hbar \frac{d}{dx} \right) \Psi(x,t) dx \tag{3.41}$$

$$< x > (t) \;\; = \;\; \int_{-\infty}^{+\infty} \Psi^*(x,t) x \Psi(x,t) dx \tag{3.42}$$

The analogy with 2D real vector linear algebra is as follows:

1. $\left(-i\hbar \frac{d}{dx} \right) \equiv \hat{p}$ is the momentum operator in the 'position basis'

2. $x = \hat{x}$ is the position operator in the 'position basis'

3. $\Psi(x,t)$ is the column coordinate vector of the particle's immutable state at time t, in the 'position basis'

4. $\Psi^*(x,t)$ is the complex conjugate of the transpose of the column coordinate vector of the particle's immutable state at time t, in the 'position basis'

5. the analogy to performing a dot or inner product by multiplying like-components of each of the 'output' or 'transposed' row coordinate vectors and the 'input' column coordinate vectors, and summing over all possible coordinates (East, North), is the integration over all x values, with the only change being that the transpose now has to also include a complex conjugation because the wavefunctions are complex

This is alot to digest! Lets start the process by noting that one of the biggest differences: instead of specifying 2D real vectors using just two coordinates (two basis vectors) with corresponding real dot products or projections, in quantum mechanics the particle's state includes an infinite number of complex coordinates, one for each value of x. The resulting "space" that is spanned by allowing for all possible square integrable complex functions of x is referred to as the Hilbert space of a quantum particle (the analogy of the "space" that included all possible 2D real vectors). Note that the "vectors" representing the immutable state of the particle are infinite in dimension. More subtlety, each value of x is the equivalent to a distinct unit basis vector! (one may or may not find it handy to think of position space x being discretized, in which case the state's vector in the position basis is still infinite, but discrete). This issue will be covered in more detail after discussing the operators.

The "position operator", rather than "rotate by angle θ", is "multiply by position x". If the discretized version of space is adopted, then one can think of $\Psi(x,t)$ as an infinite dimesion column vector):

$$\overline{\Psi(x,t)} = \begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ \Psi(x_n,t) \\ \Psi(x_n+\Delta x,t) \\ \Psi(x_n+2\Delta x,t) \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \tag{3.43}$$

where space is discretized in increments of Δx. In this case, the result of the position operator acting on this state is simply represented by

$$\overline{x\Psi(x,t)} = \begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ x_n\Psi(x_n,t) \\ (x_n+\Delta x)\Psi(x_n+\Delta x,t) \\ (x_n+2\Delta x)\Psi(x_n+2\Delta x,t) \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ x_n\Psi(x_n,t) \\ x_{n+1}\Psi(x_{n+},t) \\ x_{n+2}\Psi(x_{n+2},t) \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} . \tag{3.44}$$

Using this vector notation, the expectation value of measurements of the particle position can be expressed as,

$$<x>(t) = \int_{-\infty}^{+\infty} \Psi^*(x,t)x\Psi(x,t)dx \equiv \overline{\Psi(x,t)}^\dagger \cdot \overline{x\Psi(x,t)}, \tag{3.45}$$

where the superscript $\dagger$ refers to the transposition and complex conjugation of the vector (its adjoint, or Hermitian conjugate). In this somewhat unusual notation, one must keep

track of the fact that the time t is fixed, and it is the variable x with respect to which the dot product refers. This notation won't be used frequently, but it may be useful to fall back on if one gets confused while mastering these abstract concepts.

In a similar vein, it should be clear that to a good approximation for very small Δx, the momentum operator acts on $\Psi(x, t)$ to produce;

$$\overline{\left(-i\hbar\frac{d}{dx}\right)\Psi(x,t)} = \begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ -i\hbar\left(\Psi(x_{n+1}, t) - \Psi(x_{n-1}, t)\right)/2\Delta x \\ -i\hbar\left(\Psi(x_{n+2}, t) - \Psi(x_n, t)\right)/2\Delta x \\ -i\hbar\left(\Psi(x_{n+3}, t) - \Psi(x_{n+1}, t)\right)/2\Delta x \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}, \tag{3.46}$$

Note that while the familiar equation for the expectation value of some operator associated with a physical quantity was used above to segue from 2D vector space linear algebra to the linear algebra of quantum Hilbert space, it is more general to consider inner products of the form

$$P_{1,2}(t) \quad = \quad \int_{-\infty}^{+\infty} \Psi_1^*(x, t)\left(-i\hbar\frac{d}{dx}\right)\Psi_2(x, t)dx \equiv \overline{\Psi_1(x, t)}^\dagger \cdot \overline{\left(-i\hbar\frac{d}{dx}\right)\Psi_2(x, t)} \tag{3.47}$$

$$X_{1,2}(t) \quad = \quad \int_{-\infty}^{+\infty} \Psi_1^*(x, t)x\Psi_2(x, t)dx \equiv \overline{\Psi_1(x, t)}^\dagger \cdot \overline{x\Psi_2(x, t)} \tag{3.48}$$

that are perfectly well defined for any valid functions $\Psi_1(x, t), \Psi_2(x, t)$ in Hilbert space. These equations can be case in bra ket notation in direct analogy with the 2D real vector space equations, with the understanding that the bra now includes a complex conjugation along with the transposition of the corresponding ket. Thus

$$P_{1,2}(t) \quad = \quad \langle\Psi_1(x, t)|\hat{p}|\Psi_2(x, t)\rangle \tag{3.49}$$

$$X_{1,2}(t) \quad = \quad \langle\Psi_1(x, t)|\hat{x}|\Psi_2(x, t)\rangle \tag{3.50}$$

This notation begs the question: what is the quantum Hilbert space equivalent to the 2D real vector space *dual operators* corresponding to $\hat{x}$ and $\hat{p}$? Recalling that the general defining equation for the *dual operator* $\hat{O}^T$ of an operator $\hat{O}$ was

$$\langle\vec{r}\,'|\left(\hat{O}|\vec{r}\rangle\right) = \langle\left(\hat{O}^T|\vec{r}\,'\rangle\right)|\vec{r}\rangle = \langle\vec{r}\,'|\hat{O}|\vec{r}\rangle, \tag{3.51}$$

it makes sense that the defining equations for the *dual operators* of $\hat{x}$ and $\hat{p}$, denote them as $\hat{x}^\dagger$ and $\hat{p}^\dagger$, are

$$\int_{-\infty}^{+\infty} \Psi_1^*(x, t)x\Psi_2(x, t)dx \quad = \quad \int_{-\infty}^{+\infty} \Psi_1^*(x, t)\hat{x}\Psi_2(x, t)dx \tag{3.52}$$

$$= \quad \int_{-\infty}^{+\infty} \left(\hat{x}^\dagger\Psi_1(x, t)\right)^* \Psi_2(x, t)dx, \tag{3.53}$$

and,

$$\int_{-\infty}^{+\infty} \Psi_1^*(x,t) \left(-i\hbar \frac{d}{dx} \right) \Psi_2(x,t) dx \;=\; \int_{-\infty}^{+\infty} \Psi_1^*(x,t) \hat{p} \Psi_2(x,t) dx \qquad (3.54)$$

$$=\; \int_{-\infty}^{+\infty} \left(\hat{p}^\dagger \Psi_1(x,t) \right)^* \Psi_2(x,t) dx. \qquad (3.55)$$

The first of these is obviously satsified if $\hat{x}^\dagger$ is just $\hat{x}$, but it is less trivial to find $\hat{p}^\dagger$. Noting that if the wavefunctions go to zero at $x = \pm\infty$, then integration by parts yields,

$$\int_{-\infty}^{+\infty} \Psi_1^*(x,t) \frac{d}{dx} \Psi_2(x,t) dx = \int_{-\infty}^{+\infty} \left(-\frac{d}{dx} \Psi_1(x,t) \right)^* \Psi_2(x,t) dx, \qquad (3.56)$$

and therefore because of the factor of i included in the momentum operator, the complex conjugation cancels the negative sign associated with swapping the derivative operator from acting on $\Psi_2(x,t)$ to acting on $\Psi_1(x,t)$, with the result that

$$\hat{p}^\dagger = \hat{p}. \qquad (3.57)$$

The conclusion is therefore that in the cases of the position operator $\hat{x} \equiv$ "multiply by x" and the momentum operator $\hat{p} \equiv \left(-i\hbar\frac{d}{dx} \right)$, the corresponding operators $\hat{x}^\dagger, \hat{p}^\dagger$ that respectively satisfy

$$\int_{-\infty}^{+\infty} \Psi_1^*(x,t) \hat{x} \Psi_2(x,t) dx \;=\; \int_{-\infty}^{+\infty} \left(\hat{x}^\dagger \Psi_1(x,t) \right)^* \Psi_2(x,t) dx \qquad (3.58)$$

$$\int_{-\infty}^{+\infty} \Psi_1^*(x,t) \hat{p} \Psi_2(x,t) dx \;=\; \int_{-\infty}^{+\infty} \left(\hat{p}^\dagger \Psi_1(x,t) \right)^* \Psi_2(x,t) dx \qquad (3.59)$$

are just $\hat{x}^\dagger = \hat{x}$ and $\hat{p}^\dagger = \hat{p}$. In quantum Hilbert space these *dual operators* are instead referred to as *adjoint* or *Hermitian conjugate* operators.

The astute reader will likely now be asking why has there yet been no mention of the matrix representation of the operators "multiply by x", or $\left(-i\hbar\frac{d}{dx} \right)$? Indeed, in Eqn. 3.48, the differential form of the operators was applied to the wavefunctions before converting the "operated on" wavefunctions into discretized vectors as in Eqn. 3.44, 3.46, rather than using the matrix representation of the operators and the vector representation of the state itself. i.e. instead of writing them as,

$$P_{1,2}(t) \;=\; \int_{-\infty}^{+\infty} \Psi_1^*(x,t) \left(-i\hbar \frac{d}{dx} \right) \Psi_2(x,t) dx \equiv \overline{\Psi_1(x,t)}^\dagger \, \overline{\hat{p}} \, \overline{\Psi_2(x,t)} \qquad (3.60)$$

$$X_{1,2}(t) \;=\; \int_{-\infty}^{+\infty} \Psi_1^*(x,t) x \Psi_2(x,t) dx \equiv \overline{\Psi_1(x,t)}^\dagger \, \overline{\hat{x}} \, \overline{\Psi_2(x,t)}. \qquad (3.61)$$

The generic answer to this very valid question is that the matrix representations of the differential operators in the "position basis" can and will be found eventually, but they

are not particularly useful, in practice. However, it is worth noting that when the matrix representations of the operators are found, in whatever basis, their adjoint operators, as generally defined above, have a matrix representation given by the complex conjugate of the transpose (i.e. the adjoint) of the matrix representation of the corresponding operator, hence the name.

3.4 Hermitian Operators

It turns out that the operator associated with it any physically observable quantity, call it $\hat{Q}$ always has an adjoint, or Hermitian conjugate operator, that is equal to itself ($\hat{Q}^\dagger = \hat{Q}$, just as was shown above for the position and momentum operators). Such operators are referred to as Hermitian, and there are a number of very useful characteristics of Hermitian operators that follow immediately from this defining attribute. These include:

1. **Any operator that has a real expectation value for any valid wavefunction in Hilbert space must be Hermitian.**

2. **States that have zero variance for some Hermitian operator must be eigen states of that operator with an expectation value equal to the eigen value. Note that this does not actually imply that all eigen values of all Hermitian operators are real!**

3. **If a Hermitian operator has a discrete spectrum of eigen values (possibly with some degeneracy that will be dealt with later), then**

   (a) **the eigen values are real,**

   (b) **the associated set of eigen functions are orthonormal,**

   (c) **this orthonormal set of eigen functions are also complete (axiom),**

These properties mean that any valid wavefunction, $\Psi(x, t)$ at some specific time, can be expanded as a sum over the eigen functions of *any Hermitian operator* (recall that when the eigen functions of the Hamiltonian operator for a given potential are used for this expansion, each of the coefficients has a simple harmonic time dependence, and represent the stationary states of the system). More generally, when the wavefunction at some time t_0 is expanded in terms of the eigen states of some arbitrary Hermitian operator, there is no guarantee that the expansion coefficients will evolve harmonically as time progresses. Nevertheless, considering the expansion of the wavefunction in the bases of different Hermitian operators offers many useful insights as to the properties of the state.

Let an arbitrary quantum state of a particle be represented by $|S(t)\rangle$. Drawing on the 2D real vector space analogy, the state can be expanded in terms of the complete set of eigen states, $\{|Q_n\rangle\}$ of any Hermitian operator $\hat{Q}$, as,

$$|S(t)\rangle = \left(\sum_n |Q_n\rangle\langle Q_n| \right) |S(t)\rangle = \sum_n \langle Q_n|S(t)\rangle|Q_n\rangle. \tag{3.62}$$

Note that there is no reference to the position variable x in this equation. This emphasizes, yet again, the distinction between the immutable state of the particle, and its *representation* in any one of many possible bases (recall the immutable 2D real vector, versus its column vector representation in any arbitrary orthogonal basis $\hat{\alpha}_1, \hat{\alpha}_2$). Not only is $|S(t)\rangle$ immutable, but so are the eigen states $\{|Q_n\rangle\}$ of the Hermitian operator. The inner products $\{\langle Q_n|S(t)\rangle\}$ are also immutable. Much useful "quantum mechanics" can be done leaving equations in these immutable forms, without reference to any basis, but sometimes it is necessary to actually perform the inner product calculations. To do this, one can choose any basis in which to expand $\langle Q_n|$ and $|S(t)\rangle$, but just as in the 2D real vector space situation, there is usually a *natural* or *goto* basis used for these calculations (the $\hat{i}, \hat{j}$ basis in 2D real vector space). In quantum mechanics the goto basis is typically the position space basis, and the expansion coefficients of a state $|S(t)\rangle$ in the position basis are denoted $\Psi_{S(t)}(x, t)$, where again, to be explicit, this is to be considered a function of x at some time t in the context of it representing a set of expansion coefficients. Since the position x of the particle is a physical observable, the eigen states of the $\hat{x}$ operator can be used in the general expansion equation above to express $|S(t)\rangle$ in terms of $\Psi_{S(t)}(x, t)$;

$$|S(t)\rangle = \left(\sum_n |x_n\rangle\langle x_n| \right) |S(t)\rangle = \sum_n \Psi_{S(t)}(x_n, t)|x_n\rangle. \tag{3.63}$$

Here we have adopted the discretized vision of space as earlier, to assist conceptually. Later this will be generalized to a continuous eigen spectrum. This equation has to be augmented with the defining equation for $\{|x_n\rangle\}$, which is simply,

$$\hat{x}|x_n\rangle = x_n|x_n\rangle, \tag{3.64}$$

where $\{x_n\}$ are a set of real numbers (eigen values), one for each of the discretized positions in 1D space, and $\{|x_n\rangle\}$ are the corresponding immutable eigen states of the Hermitian position operator $\hat{x}$. Thus

$$\Psi_{S(t)}(x_n, t) = \langle x_n|S(t)\rangle. \tag{3.65}$$

Alot has been introduced in the last paragraph! The basis-independent particle state at some time t, $|S(t)\rangle$, is broken down, or expanded in terms of its components along each of the basis states *of the position basis*, by inserting the expansion of the unity operator, $(\sum_n |x_n\rangle\langle x_n|)$ in the position basis (that simply reflects its complete, orthonormal nature), and the resulting projections of any given eigen state of the position vector onto $|S(t)\rangle$, $\langle x_n|S(t)\rangle$ define, or are labelled as the wavefunction, or more explicitly, the position space wavefunction, $\Psi_{S(t)}(x_n, t)$.

If the eigen states of the momentum operator are used to expand the state $|S(t)\rangle$, then the unity operator expressed in terms of the momentum basis states is used, and one gets,

$$|S(t)\rangle = \left(\sum_n |p_n\rangle\langle p_n| \right) |S(t)\rangle = \sum_n \Phi_{S(t)}(p_n, t)|x_n\rangle. \tag{3.66}$$

Now there is a different set of expansion coefficients, $\Phi_{S(t)}(p_n, t)$ because the basis used for the expansion has changed (directly analogous to when the coordinates of a 2D real vector were changed when the $\hat{\alpha}_1, \hat{\alpha}_2$ basis was used to describe it rather than the conventional $\hat{i}, \hat{j}$ basis). The collection of expansion coefficients of $|S(t)\rangle$ in terms of the set of momentum eigen functions $\{|p_n\rangle\}$ defined by $\hat{p}|p_n\rangle = p_n|p_n\rangle$, is referred to as the momentum space wave-function. *It is no more or less a description of the particle's state $|S(t)\rangle$ than is $\Psi_{S(t)}(x_n, t)$.*