Notes for Lecture 1

Formalism of Quantum Mechanics

Here in the first lecture we briefly discuss what lies ahead and summarize some essential elements of Quantum Mechanics from the formalism point of view.

1.1 QM I and QM II

As you are taking the second installment of the introductory quantum mechanics, you must have enjoyed learning quantum mechanics so far. This is, of course, very nice, and, so, Kudos to you! Quantum mechanics is a relatively new discipline, it is one of the most fundamental of all physics, and there is much to be excited about in it.

In the first installment (let us call it QM I), you typically learn the basic problems of Quantum Mechanics, such as simple potential well problems, the simple harmonic oscillator problem, and the hydrogenic atom problem. Also, importantly, you learn the basic formalism: the Schrödinger equation, the uncertainty principle, the measurement hypothesis, the mathematics of the Hilbert space, identical particles, the exchange symmetry, and the Dirac notation.

Now, this course is the second installment (let us call it QM II). This means that it is somewhat more advanced QM. It also means that it is somewhat less pretty QM. What do I mean by this? What you learn in QM I is very pretty and impressive, but there is a certain distance between it and the reality. Most problems that you learn in QM I are simple and exactly solvable. They are very nice and very essential things to learn. They are something like museum pieces. Essential as they are, however, the problems that you learn in QM I often do not describe real problems

very well. They are hardly enough to let you explain real phenomena. To account for the real experimental data, one has to do some more. This "doing more" is what QM II is mostly about. So, here in QM II, we learn practical techniques such as the perturbation theory and the variational principle.

At first, these approximation methods may look like lesser principles of secondary importance. However, note that such methods are the only ways that we can use to confront our theory with our experiment. And it goes without saying that such a confrontation is what makes science progress. In this sense, the approximation methods are as essential parts of physics as any other methods or principles.

One can generally say that the reason why one needs approximation methods in the first place is because there are interactions. No particle is free. Or, if there was a truly free particle, we would not know about it. Learning how to deal with a myriad of interactions by being able to identify the most important interaction, or a few most important interactions, is a big part of doing quantum mechanics, or physics in general.

Lastly, one might note the following. Any modern physics student starts out the physics learning by first learning classical mechanics, starting from the venerable notion of a "particle." A "classical particle" is, among other things, a notion of something that is unquestionably in existence for a certain period of time, assuming of course that the environment is reasonably mild. So, in classical mechanics, a figure skater is a particle, a tennis ball is a particle, as is the Earth or the Sun. In quantum mechanics, we look deeper and deeper into each of these particles. What do we discover? We discover that a "classical particle" consists of a whole bunch of more elementary "quantum particles" – electrons, protons, neutrons, etc. We also discover that these elementary quantum particles occupy a completely negligible volume of the original macroscopic particle. That is, in some sense, a classical particle is almost completely empty inside! Last but not least, what we discover is that there is no way in the world to distinguish between electrons that make up one classical particle, say a knife, and electrons that make up another classical particle, say an apple. Then, one may ask: how is it that we cannot go through walls, a blade can cut an apple, and an apple cannot cut a blade? The answer is interactions among a large number of particles. The notion of a classical particle as a sure, sometimes self-aware, thing can arise only through interactions of very many particles. Also, the notion of a distinguishable particle and its hardness are due to those interactions. This notion of "a substantive particle emerging out of interacting trivial particles" is quite general and goes beyond the current discussion: the Higgs mechanism is an example of the same notion. In QM II, our job is to start our journey to deal with these interactions, a fundamental ingredient of Nature.

1.2 QM Formalism

Here, let us review the formalism of Quantum Mechanics. We also establish important Dirac notation, as we do so. We will use the Dirac notation throughout this course.

In reading this note, keep in mind that it is kind of hard to know the whole QM formalism clearly even when you can do many certain QM problems. Do not be disappointed if you find things that you feel you should know by now but you don't. This is normal. Doing things and knowing things are different and you need to have a bit of both. Here, this review is a grand reminder, presented to keep you informed of the abstract structure of QM, before we delve into actual computations. When you come back to this note later, things may look much clearer. You will understand QM much better. And then ... this will in turn help you do QM problems even more easily!

1.2.1 Kinematics

Let us lay out some basic kinematical ingredients necessary to describe QM. In some sense, the kinematical ingredients described here are no more complicated than those encountered in Classical Mechanics: the starting point is a vector space where angles and lengths can be measured and the calculus can be performed. However, in QM, the vector space is a complex vector space and its dimension is usually very large (infinite), since we are concerned with a state vector, not a position vector or a velocity vector as in a single particle dynamics of Classical Mechanics[1].

State Vector

A quantum mechanical state is denoted as $|\Psi\rangle$ (a ket vector) or $\langle\Psi|$ (a bra vector).

An example of a state is a neutrino zooming in space, a hydrogen atom, a cup filled with water, or you.

A state is a physical entity. Mathematically, a physical state corresponds to a non-zero vector[2]. The two vectors $\langle\Psi|$ and $|\Psi\rangle$ correspond to the same state, assuming that the symbol Ψ denotes a certain state that we are interested in. The two ways of

[1] However, if a wave phenomenon is described in Classical Mechanics, then the vector space necessary for that is very much like that necessary for a QM problem of a single particle.

[2] A zero vector is truly nothing, not a vacuum state.

referring to the same state called "duality." I.e., $|\Psi\rangle$ is referred to as the dual of $\langle\Psi|$ and vice versa.

For two states, $|\alpha\rangle$ and $|\beta\rangle$, the product $\langle\alpha|\,\beta\rangle$ is a complex number. This is the so-called *inner product* of two vectors.

When states/vectors are *represented* (to be explained in Lecture 2), ket vectors correspond to column vectors, and bra vectors correspond to row vectors.

The following relation is valid.

$$\langle\alpha| \;=\; |\alpha\rangle^\dagger \tag{1.1}$$

$$|\alpha\rangle \;=\; \langle\alpha|^\dagger \tag{1.2}$$

The notation † means the "Hermitian conjugation." Another name is "adjoint" operation. It means the transpose operation followed by the complex conjugation of all elements. The above relation means that the bra vector for an arbitrary state α is the Hermitian conjugate of the ket vector for the same state, and vice versa.

Norm, Hilbert space

The norm of a vector $|\alpha\rangle$, can be denoted as $N(\alpha)$, and it is given by

$$N(\alpha) \;=\; \sqrt{|\langle\alpha|\,\alpha\rangle|} \tag{1.3}$$

Any physical state has a finite norm. By *normalizing*, any physical state can be made to have a unit norm. This normalization convention is usually applied to any physical state. The reason to choose this convention will become clear when we discuss probabilities below.

A vector space where the inner product is well-defined, the norm of any vector is finite, and the limit of a converging sequence of vectors is also a vector (i.e., a "complete" metric space) is called the "Hilbert space." Put more physically, the Hilbert space means a vector space where we can measure lengths and angles, and also do calculus.

For a given quantum mechanical problem, a Hilbert space is defined by all possible physical states, up to an arbitrary complex normalization constant, and the null vector (i.e. the origin, or the zero vector, of the vector space). So, in a Hilbert space, any vector except the origin corresponds to a physical state, up to a normalization constant.

Often, it is helpful to extend the Hilbert space to include some states that do not have a finite norm. Then we get a space called a "rigged," or an equipped, Hilbert

space. Such a Hilbert space allows to include vectors that have diverging norms, but diverging in the manner of the Dirac delta function. Such rigged Hilbert space can be considered a purely mathematical invention, but it is an extremely useful one. Wave functions such as $\frac{1}{\sqrt{2\pi}}\exp(ikx)$ or $\delta(x-x_0)$ can exist only in such rigged Hilbert space, while in reality there is no way to prepare a truly omnipresent plane wave or a truly point-localized wave (equivalent to a truly omnipresent plane wave in the *momentum space*). In other words, the distinction between the rigged Hilbert space and the Hilbert space is necessary only when a truly infinite space-time must be considered. For finite, but large, space-time, the Hilbert space is enough to consider.

With this in mind, from now on, we will simply refer to our state space for any given problem as "Hilbert space" even when extended states are thrown in.

The dimensionality of the Hilbert space for any given problem is defined by the nature of the problem. Consider a particle in a box in one dimension. I.e., let us say that a particle can exist from $x = 0$ to $x = L$, with L some finite number. What is the dimension of the Hilbert space? Since the particle can exist at any point, and since we have an infinite number of points, it is infinite. Likewise, the Hilbert space of the hydrogen problem is infinite. However, some problems have finite dimensions. For instance, if one is measuring the spin quantum number of a beam of Ag atoms, and if one is interested in the spin quantum number alone, then the dimension of the Hilbert space is 2, since it happens that a Ag atom has a half spin.

Operator

In quantum mechanics, the term *operator* is used specifically for a linear transformation from a Hilbert space to the *same* Hilbert space. So, an *operator* corresponds to a *square* matrix.

A quantum mechanical operator is denoted as $\hat{O}$. In this course, I will try my best to put a "hat" on top of an operator, to explicitly distinguish it from a mere number[3].

Suppose an operator $\hat{O}$ has the property $\hat{O}|\alpha_1\rangle = |\alpha_2\rangle$ and $\hat{O}|\beta_1\rangle = |\beta_2\rangle$. Then, for any numbers C_1 and C_2, we have

$$\hat{O}\left(C_1|\alpha_1\rangle + C_2|\beta_1\rangle\right) = C_1\hat{O}|\alpha_1\rangle + C_2\hat{O}|\beta_1\rangle = C_1|\alpha_2\rangle + C_2|\beta_2\rangle \tag{1.4}$$

due to the operator being defined to mean a *linear* transformation. This superposition principle is why quantum mechanics is often called *wave mechanics*.

[3]However, outside an introductory quantum mechanics course, one often forgoes the hat notation, when the context makes it clear whether we are discussing an operator or a number.

Hermitian operator

A Hermitian operator $\hat{h}$ is an operator that satisfies the following

$$\hat{h}^\dagger = \hat{h} \tag{1.5}$$

where the Hermitian conjugation (†) operation is, as before, the transpose operation followed by the complex conjugation of all elements of the matrix. Other equivalent names for the Hermitian operator are the self-adjoint operator or the real operator. The reason why a Hermitian operator is also called a *real* operator is because all eigenvalues of a Hermitian operator are real.

A Hermitian operator is always diagonalizable, and its eigenvectors can be taken to be (Dirac-)orthonormal (see sub-section "Natural basis" below).

Unitary operator

Another class of important operators are unitary operators. A unitary operator is an operator that conserves the norm of any vector. Starting from this definition, one can prove that the following three conditions are all equivalent to one another[4].

$$\left\langle \hat{U}\alpha \middle| \hat{U}\alpha \right\rangle = \langle \alpha | \alpha \rangle \qquad \text{for any } |\alpha\rangle \tag{1.6}$$

$$\left\langle \hat{U}\alpha \middle| \hat{U}\beta \right\rangle = \langle \alpha | \beta \rangle \qquad \text{for any } |\alpha\rangle \text{ and } |\beta\rangle \tag{1.7}$$

$$\hat{U}^\dagger \hat{U} = \hat{U}\hat{U}^\dagger = 1 \qquad \text{i.e., } \hat{U}^\dagger = \hat{U}^{-1} \tag{1.8}$$

where 1 here means the unit matrix, i.e. the identity matrix [5]

In quantum mechanics the vector space that we consider is a complex vector space. If the vector space were real, then an inner-product conserving operator would be called an orthogonal operator.

Here, $\left|\hat{U}\alpha\right\rangle \equiv \hat{U}|\alpha\rangle$. The proof of the equivalence of these relations is left for exercise[6], for the most part. Here, we show how to go from the second statement to the third statement. Using, at the third step, the well-known property from Linear

[4]While we use the term "operator" to include only linear transformations, in a more advanced QM course, it is sometimes necessary to generalize this notion a bit to include an *anti-linear* transformation (see footnote 7). In that case, an anti-unitary operator is also norm-conserving.

[5]Any number added to, subtracted from, or compared with, an operator denotes an identity operator/matrix times that number, by the same token.

[6]The proof can be found in `http://griffin.ucsc.edu/teaching/08Q2-139A/download/L8%20-%20Formalism.pdf`.

Algebra that $(AB)^\dagger = B^\dagger A^\dagger$, we see that the above definition implies

$$\left\langle \hat{U}\alpha \middle| \hat{U}\beta \right\rangle = \left(\left| \hat{U}\alpha \right\rangle \right)^\dagger \left| \hat{U}\beta \right\rangle = \left(\hat{U} \left| \alpha \right\rangle \right)^\dagger \hat{U} \left| \beta \right\rangle = \left\langle \alpha \right| \hat{U}^\dagger \hat{U} \left| \beta \right\rangle = \left\langle \alpha \middle| \beta \right\rangle$$

for any two states $|\alpha\rangle$, $|\beta\rangle$. The last two terms can be equal to each other for arbitrary choices of $|\alpha\rangle$ and $|\beta\rangle$ if and only if $\hat{U}^\dagger \hat{U} = 1$. This means that a unitary operator is an operator, whose adjoint is its inverse.

Rotations, reflections, and inversion are some examples of a unitary operation. This may not be surprising. However, in quantum mechanics, the time evolution operator and the translation operator are also unitary operators[7].

A unitary operator is also always diagonalizable with (Dirac-)orthonormal eigenvectors. Eigenvalues of a unitary operator are modulus 1 complex numbers, $\exp(i\theta)$ with a real number θ.

Observable

An observable is a dynamical variable that can be measured in an experiment, or in an "event." Examples are mass, charge, particle number, position, area, volume, momentum, angular momentum, energy, etc. In quantum mechanics, time is not an observable, since it is not considered a dynamical variable; it is merely a parameter.

A mathematical object that corresponds to an observable is a *Hermitian/real/self-adjoint operator*. Thus, in QM, any observable corresponds to a Hermitian *matrix*, which is necessarily a *square* matrix. Please re-read the previous paragraph of this section with this fact in mind.

However, one should keep in mind that it takes a somewhat higher level of QM to appreciate why mass, charge, particle number etc. must correspond to matrices. In a single particle QM, which we are concerned with here mostly, position, momentum, angular momentum, energy etc. are important operators to consider.

[7] There are also certain quantum-mechanical operators, such as the time reversal operator, that are "anti-unitary." We may not deal with them in this course, but you will learn them in more advanced courses. An operator $\hat{\theta}$ is anti-unitary if $\left\langle \hat{\theta}\alpha \middle| \hat{\theta}\beta \right\rangle = \left\langle \beta \middle| \alpha \right\rangle$ and θ is anti-linear, i.e. $\hat{\theta}\left(c_1 \left|\alpha\right\rangle + c_2 \left|\beta\right\rangle\right) = c_1^* \hat{\theta} \left|\alpha\right\rangle + c_2^* \hat{\theta} \left|\beta\right\rangle$ for any states α, β and for any complex numbers c_1 and c_2. So, an anti-unitary operator is *not* an operator within our current (narrow) definition, where an operator is required to be linear: $\hat{L}\left(c_1 \left|\alpha\right\rangle + c_2 \left|\beta\right\rangle\right) = c_1 \hat{L} \left|\alpha\right\rangle + c_2 \hat{L} \left|\beta\right\rangle$.

Natural basis

While the choice of basis for a given vector space is completely arbitrary in principle, we tend to choose the most useful ones in practice.

Let us consider a Hermitian operator. Consider its eigenvectors $\{e_i\}$. In this notation, we are assuming that eigenvalues are discrete $i = 0, 1, 2, ...$ It is also possible that some or all of the eigenvalues form a continuum. In that case the subscript i is unnecessary, and e can be taken to be a continuous variable (or i can be re-interpreted as a continuous index)[8].

Eigenvectors of a Hermitian operator are (Dirac)-orthonormal, which means that

$$\langle e_i | e_j \rangle = \delta_{i,j} \qquad \text{discrete eigenvalue spectrum} \tag{1.9}$$

$$\langle e | e' \rangle = \delta(e - e') \qquad \text{continuous eigenvalue spectrum} \tag{1.10}$$

Also, they span the Hilbert space. That is,

$$|\alpha\rangle = \sum_i c_i |e_i\rangle + \int de\, c(e) |e\rangle \tag{1.11}$$

where c_i's and $c(e)$ are determined uniquely for any state $|\alpha\rangle$. Here, the most general case where the eigenvalue spectrum is partly discrete and partly continuous is considered. In many cases, however, the eigenvalue spectrum is either completely discrete or completely continuous – in those cases, only the first term or only the second term suffices.

Generally, a set of basis satisfying the above properties are called the *natural basis*. There is hardly any reason why a non-natural basis should be considered for any problem.

A natural basis has the following property (**the resolution of identity**).

$$1 = \sum_i |e_i\rangle \langle e_i| \qquad \text{bound basis states; discrete spectrum} \tag{1.12}$$

$$1 = \int de\, |e\rangle \langle e| \qquad \text{extended basis states; continuous spectrum} \tag{1.13}$$

$$1 = \sum_i |e_i\rangle \langle e_i| + \int de\, |e\rangle \langle e| \qquad \text{mixed case; Note that } \langle e | e_i \rangle = 0. \tag{1.14}$$

[8] In this sub-section, e simply means "eigen-value." It has nothing to do with the number $\exp(1) \approx 2.718$.

Compatible operators

Consider two operators, $\hat{A}$ and $\hat{B}$, that are separately diagonalizable. If the **commutator** is zero ($[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} = 0$), i.e. if the two operators commute, then these two operators are said to be "compatible," and it is possible to find eigenvectors that can diagonalize the two operators simultaneously.

1.2.2 Dynamics

Now that we are through defining vectors and matrices, let us imbue them with meanings – mechanical meanings.

Measurement

The word "measurement" is arguably not the best word to use here, but it is widely used in the literature. Some prefer the word "event."

While in most situations, the word "measurement" is not too misleading, the following two points must be kept in mind about the measurement.

1. The measurement need *not* involve any human involvement, despite the mundane connotation of the word.

2. A measurement does need to involve a "piece of equipment" – a classical mechanical object that can clearly record a certain value of the measured quantity. This equipment need not be man-made.

So, a quantum mechanical measurement means that a quantum mechanical system interacts with equipment (a classical system) and makes the equipment show a certain event (like a moving needle, a photocell lighting up, a ball rolling down in a certain direction, etc.), whether or not a human being is there to observe this event.

Let us consider a measurement of an observable $\hat{A}$ on a state $|\Psi\rangle$. The following constitute the QM measurement hypothesis.

1. Equipment records *one of the eigenvalues* of the observable $\hat{A}$. Note that this value must be real due to the Hermitian matrix nature of any observable.

2. After the measurement the state of the quantum system is given by $|A\rangle$, the eigenvector corresponding to the eigenvalue A measured. The transition from $|\Psi\rangle$ to $|A\rangle$ is often referred to as a wave function collapse.

3. The probability that this measurement will yield value A is given by $|\langle A|\Psi\rangle|^2$. This statement is unambiguous for interpretation, if the set of eigenvalues $\{A\}$ form a discrete set. If the eigenvalues $\{A\}$ form a continuous set, then this statement can be made a bit more transparent: the probability that a measurement will yield a value between A and $A + dA$ (or a value that falls into any small interval dA around value A) is given by $|\langle A|\Psi\rangle|^2 dA$.

The last statement assumes that $|\Psi\rangle$ is normalized. Only then, the sum/integral of all probabilities is 1, which must be the case. This can be proven using the resolution of identity (see sub-section "Natural basis" above).

Schrödinger equation

The measurement hypothesis and the time evolution dictated by the Schrödinger equation form the fundamental dynamical principles of quantum mechanics. The Schrödinger equation is

$$i\hbar\frac{d}{dt}|\Psi\rangle = \hat{H}|\Psi\rangle \tag{1.15}$$

Here the left hand side gives information the rate at which the state changes in time, and the right hand side is simply a product of a matrix with a vector. Thus, the physical content of this equation is that once you know the state at a given time, then you know what is going to happen to it at a later time or what happened earlier. This determinism of Schrödinger equation is the same level of determinism of Newton's second law.

Note that the above Schrödinger equation has a total time derivative on the left hand side. This is because the state $|\Psi\rangle$ does not depend on any other variables than t (this point will be explained more fully in Lecture 2). So, the above equation can be re-written in the following form.

$$\begin{aligned}
|\Psi(t+dt)\rangle &= |\Psi(t)\rangle + dt\cdot\frac{d}{dt}|\Psi(t)\rangle \\
&= |\Psi(t)\rangle + dt\cdot\frac{\hat{H}}{i\hbar}|\Psi(t)\rangle \qquad \text{using Eq. 1.15} \\
&= \left(1 - i\,\frac{\hat{H}}{\hbar}\,dt\right)|\Psi(t)\rangle
\end{aligned} \tag{1.16}$$

This last expression makes the meaning of Schrödinger equation truly transparent, as follows.

$$|\Psi(t+dt)\rangle \;=\; \hat{\mathcal{U}}(dt)\,|\Psi(t)\rangle \tag{1.17}$$

$$\hat{\mathcal{U}}(dt) \;=\; 1 - i\,\frac{\hat{H}}{\hbar}\,dt \qquad\qquad \textbf{time evolution operator} \tag{1.18}$$

Eq. 1.18 means that the Hamiltonian drives non-trivial time evolution (put mundanely, you die when there is no energy): in physics terms, $\hat{H}$ is the **generator** for time evolution.

Eq. 1.18 is valid in general, i.e. for *any* $\hat{H}$, e.g. even when $\hat{H}$ contains an explicit time dependence, as would be the case if it describes an open system influenced by the environment.

If $\hat{H}$ does not explicitly depend on time, then Eq. 1.18 can be integrated trivially. In *this* case,

$$\hat{\mathcal{U}}(\Delta t) \;=\; \left[\hat{\mathcal{U}}(dt)\right]^{\Delta t/dt} \tag{1.19}$$

$$\;=\; \left(1 - i\,\frac{\hat{H}}{\hbar}\,dt\right)^{\Delta t/dt} \tag{1.20}$$

$$\;=\; \left[(1+\delta)^{1/\delta}\right]^{\Delta t\,\delta/dt} \qquad\qquad \delta \equiv -i\hat{H}dt/\hbar \tag{1.21}$$

$$\;=\; \exp\left(-i\frac{\hat{H}\Delta t}{\hbar}\right) \qquad\qquad \lim_{\delta\to 0}(1+\delta)^{1/\delta} = e, \quad \Delta t\,\delta/dt = -i\hat{H}\Delta t/\hbar \tag{1.22}$$

giving

$$|\Psi(t+\Delta t)\rangle \;=\; \hat{\mathcal{U}}(\Delta t)\,|\Psi(t)\rangle \tag{1.23}$$

or, changing the time-axis notation slightly,

$$|\Psi(t)\rangle \;=\; \hat{\mathcal{U}}(t)\,|\Psi(t=0)\rangle \;=\; \exp\left(-i\frac{\hat{H}t}{\hbar}\right)|\Psi(t=0)\rangle \tag{1.24}$$

This result can be derived, just as easily, from directly integrating Eq. 1.15. In any case, this result should look sensible to those with some experience in doing QM problems, for the following reason. It shows that when $|\Psi(t=0)\rangle$ is expressed as a linear combination of $\hat{H}$ eigenstates $\{|E_n\rangle\}$, where E_n is the n-th energy eigenvalue, the time evolution of each eigenstate is simply given by the phase factor $\exp(-iE_nt/\hbar)$. Namely, if

$$|\Psi(t=0)\rangle \;=\; \sum_n C_n\,|E_n\rangle \tag{1.25}$$

then

$$|\Psi(t)\rangle \;=\; \sum_n C_n \hat{\mathcal{U}} |E_n\rangle \;=\; C_n e^{-iE_n t/\hbar} |E_n\rangle \tag{1.26}$$

where the $\sum_n$ must be replaced by $\int dE$ for any part of E_n that forms a continuous spectrum, not a discrete spectrum.

When *represented* in real space (to be derived in Lecture 2), the Schrödinger equation, Eq. 1.15, becomes the following more familiar equation, involving a partial time derivative

$$i\hbar \frac{\partial}{\partial t} \Psi(\vec{x}, t) \;=\; H\left(\vec{x}, \vec{p}, t\right)\big|_{\vec{p}=-i\hbar\frac{\partial}{\partial \vec{x}}} \Psi(\vec{x}, t) \tag{1.27}$$

where $\Psi(\vec{x}, t)$ and $H\left(\vec{x}, \vec{p}, t\right)\big|_{\vec{p}=-i\hbar\frac{\partial}{\partial \vec{x}}}$ are now both *representations* of the state $|\Psi\rangle$ and $\hat{H}$ respectively, in real space.