Notes for Lecture 2

Formalism, cont.

2.1 QM Formalism

We discuss some more of QM formalism, with emphasis on representation and symmetry/conservation.

2.1.1 Representation

The advantage of using the Dirac notation can be realized through the proper appreciation of the concept of the "representation" of a vector. Here, let me discuss this important point a bit.

The concept of representation is not necessarily something that you will have mastered by now, *in the context of QM*. Often times, this concept goes amiss during the first course of QM mechanics. However, this concept is not a difficult one by itself – it is just that the true meaning of the Dirac notation sometimes goes amiss. If this is what happened to you, then this section will correct the situation.

Let us start by asking what a vector is. A mathematician would start by listing some axiomatic properties that a certain quantity must satisfy to be qualified as a vector[1]. Essentially, they come down to these intuitive statements. (1) A quantity is a vector if it can be scaled by a number and if it can be added to another vector. (2) A null vector ("zero") must be well-defined, as must be an inverse vector. These intuitive

[1]See, e.g., Table 1 in `http://griffin.ucsc.edu/teaching/08Q2-139A/download/L8%20-%20Formalism.pdf`.

statements are indeed summarized well by the elementary visualization of "vector as arrow." An arrow can be stretched up or shrunken down (even to zero length), reversed (defining the inverse), added to another arrow (using the "parallelogram rule")[2]. This "vector as arrow" is a very effective concept, I think, because we can easily imagine and grasp an arrow in our mind.

Defining what a vector is, and visualizing it as an arrow, is great, but is not enough to do "the real work." For the real work to begin, we need to crunch numbers. To start crunching numbers, we need to first *represent* a vector/an arrow as a set of numbers. Of course, the process of representing a vector (like a position vector) as a set of numbers is such a natural one so we do it almost automatically and often optimally. We often work with $xy...$ coordinates, representing an arrow using the Cartesian basis $(\hat{x}, \hat{y}, ...)$. In this case, Cartesian coordinates correspond to the representation. Let us note that, in principle, the orientation of the Cartesian coordinate system is completely arbitrary, and so we have an infinite number of choices for this representation, while we tend to choose the most obvious one instinctively. We can also use length and angle for quantifying an arrow – in this case, we are representing an arrow using the polar basis $(\hat{r}, \hat{\theta}, ...)$ in a locally Cartesian coordinate system. In this case, polar coordinates $(r, \theta, ...)$ correspond to the representation. Here again, the representation is not unique at all, since the reference direction from which the angle is measured is completely arbitrary.

Now, equations for all physical laws are written in terms of vectors (or, more generally, tensors) for a reason. Physical laws are independent of how we take our coordinate-system/basis! This is clear, since what we experience is unique no matter what coordinate system we draw in space and time to describe it. Indeed, the plain fact is that practically all basic physical laws, such as Newton's second law ($\vec{F} = d\vec{p}/dt$) and Maxwell's equations ($\vec{\nabla} \cdot \vec{E} = \rho/\epsilon_0$) are all written in an abstract, representation-independent, way. This is the best way to do it, since physical phenomena that they describe are independent of representations. **Alas, the usual way that students first learn QM is *not* this best way.**

Fig. 2.1 is a reminder that many quantities in physics are, in fact, vectors. At an elementary level, this statement is plausible if one realizes that, if a physical quantity can be *represented* by a column of numbers (e.g., for velocity or a *function*), then that quantity is usually (while not always) a vector.

Now, let us go through each of the three examples of Fig. 2.1 with more care. Doing so will help us better grasp the concept of vector (in QM), and, in particular, clearly recognize the *difference between a vector and its representation*.

[2]You might argue that this is valid only for a *real* vector. However, for a complex vector, one can always consider the real part and the imaginary part as two separate arrows.

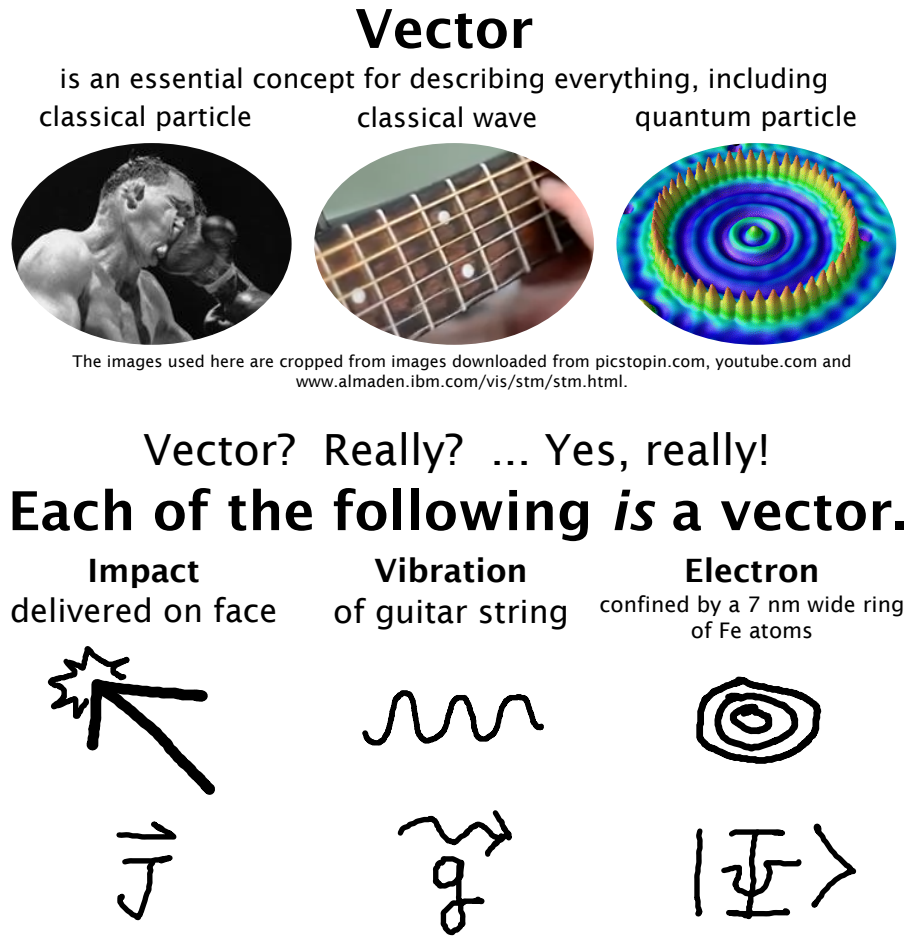

Figure 2.1: The concept of vector is essential in all physics. Here are some examples in the context of this course.

Impact on a boxer – vector for a classical particle

This is perhaps the most familiar example, if not the most obvious. In classical mechanics of a single particle, various quantities are vectors: position, momentum, force, impact, etc. Impact = force × short collision time.

Let us use the symbol $\vec{J}$ for the impact vector. How do we represent this vector? Depending on which of the three basis sets of the figure, we use, the answer is one of

$$\vec{J} \doteq \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x' \\ y' \end{pmatrix}, \text{ or } \begin{pmatrix} r \\ \theta \end{pmatrix} \tag{2.1}$$

Here, we consider a plane containing the impact vector as our vector space, and different choices of basis set for the plane result in different numbers to represent $\vec{J}$. The symbol $\doteq$ is used to mean "is represented by." In my lectures and lecture notes, I will always distinguish between vector/matrix and its representation, for the utmost clarity. The symbol $\doteqdot$ will be used for "represents," as in $\binom{x}{y} \doteqdot \vec{J}$.

Note that here we are using a *column* vector to represent $\vec{J}$. This is the correct convention to use, whenever vectors and matrices need to be considered together. I.e., it is the convention of Linear Algebra.

Where do the above representations come from? It is very helpful to *derive* them in the following way.

First of all, let us note that the following is true for 2-dimensional *complex* two dimensional vectors[3].

$$1 = \begin{pmatrix} a \\ b \end{pmatrix} \begin{pmatrix} a^* & b^* \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} \begin{pmatrix} c^* & d^* \end{pmatrix} \qquad \text{resolution of identity} \tag{2.2}$$

if $\vec{e}_1 \doteq \binom{a}{b}$ and $\vec{e}_2 \doteq \binom{c}{d}$ are orthonormal, i.e.

$$\begin{pmatrix} a^* & b^* \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = 0, \quad |a|^2 + |b|^2 = |c|^2 + |d|^2 = 1 \tag{2.3}$$

Note that, the unity on the left hand side of Eq. 2.2 means $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, following the convention defined in the last lecture (page 6). It is left for you to prove[4] Eq. 2.2 starting from Eq. 2.3.

[3] We do not need to consider complex numbers for the current classical mechanics example. But, we consider a more general complex vector space here, anyway, to appreciate the mathematics essential for QM.

[4] A brute force proof without prior knowledge is not necessarily easy, but quite doable in this simple 2×2 matrix case. To prove brute force, the fact that the determinant of the right hand side of Eq. 2.2 is 1 is useful. For a simpler, learned, proof, start by noting that a transpose of a unitary matrix is also unitary.

The above equations can be rewritten more economically as

$$1 = \sum_i \vec{e}_i \, \vec{e}_i^{\,\dagger} \qquad \text{equiv. to Eq. 2.2; same form as Eq. 1.12} \tag{2.4}$$

$$\vec{e}_i^{\,\dagger} \, \vec{e}_j = \delta_{i,j} \qquad \text{equiv. to Eq. 2.3; same form as Eq. 1.9} \tag{2.5}$$

These equations not only are extensible to any higher dimensional vector space but also have the identical form as Eqs. 1.9 and 1.12.

Now, any general vector, such as $\vec{J}$, can be expressed as

$$\vec{J} = 1 \cdot \vec{J} \tag{2.6}$$

$$= \sum_i \vec{e}_i \, \vec{e}_i^{\,\dagger} \vec{J} \qquad \text{using Eq. 2.4} \tag{2.7}$$

$$= \sum_i \vec{e}_i \left(\vec{e}_i^{\,\dagger} \vec{J} \right) \qquad \text{associativity} \tag{2.8}$$

$$= \sum_i \left(\vec{e}_i^{\,\dagger} \vec{J} \right) \vec{e}_i \qquad \vec{e}_i^{\,\dagger} \vec{J} \text{ is an inner product, and so is a number} \tag{2.9}$$

The **representation** of the vector $\vec{J}$ simply means **the coefficients** $\vec{e}_i^{\,\dagger} \vec{J}$. In other words, it is **components**. If $\vec{e}_1 = \hat{x}$ and $\vec{e}_2 = \hat{y}$, then $\hat{x}^t \vec{J} (\equiv \hat{x} \cdot \vec{J}) = J_x$ and $\hat{y}^t \vec{J} (\equiv \hat{y} \cdot \vec{J}) = J_y$, and so the representation is $\begin{pmatrix} J_x \\ J_y \end{pmatrix}$. Similarly, the other two representations of Eq. 2.1 can be understood as taking $(\vec{e}_1, \vec{e}_2) = (\hat{x}', \hat{y}')$, or $(\vec{e}_1, \vec{e}_2) = (\hat{r}, \hat{\theta})$.

Vibration of guitar string – vector for a classical wave

Now, let us consider the wave phenomenon of a guitar string, the second example in Fig. 2.1. Pick one of the guitar strings. Then, as a function of the position x along the axis of the guitar, one can record the amplitude of the wave, thereby obtaining a function. Let us call this amplitude function, $g(x)$, defined from $x = 0$ to L, the length of the string at rest.

All possible amplitude functions $g(x)$ form a vector space for the very reason that amplitude functions can be scaled up and down, they can be added and subtracted, and they include a unique null element ($g(x) = 0$ for any x)[5]. From a more physical point of view, one must realize that a function $g(x)$ will necessarily be represented by a *finite* list of numbers, e.g. a column in a spreadsheet, in any experiment, due to

[5]In this rudimentary sense, you might say even matrices can be vectors. However, notice also that the Hilbert space – which is in the end what we care about – is an inner product space, and so a vector needs to have well defined attributes such as length and angle. So, in this more detailed sense, we cannot call matrices vectors.

both *intrinsic* (different physics at smaller length scales) and extrinsic (resolution of equipment, due to finite money/time) limit of the spatial resolution. In this physical view, a finite dimensional vector is defined by $g(x_i)$, with $i = 1, ..., N$, where N is the number of sampling points. Physically plausible in every way this view may be, it is more straightforward mathematically to consider x as a *continuous* variable and consider functions $g(x)$. We will follow this mathematical convention of assuming a continuous variable x.

What is a vector quantity in this problem? And, what are its possible representations?

An answer to the second question is easy, since we already have one representation! $g(x)$. The answer to the first question is also easy, in fact. It can be written as $\tilde{\vec{g}}$, which means *the wave of the guitar string itself.*

For a given unique physical situation describable in terms of a vector, the vector is unique while its representation is not unique at all. In the previous example, $\vec{J}$ is unique, at the moment of collision, while $J_i = \vec{e}_i^{\,\dagger}\vec{J}$ is dependent on our choice of basis. In the current example, $\tilde{\vec{g}}$ is unique for the snapshot of the string vibration, but $g(x)$ is only one of the ways that that vibration can be represented.

So, what other function than $g(x)$ can be a valid representation? Here is a good example. If you are a sound engineer, you may like to find out how many overtones exist, or what wave vectors ($k = 2\pi/\lambda$) are excited. To answer the latter question, one can feed the snapshot of the string vibration to a computer, let the computer trace the shape of the string, let the computer do the fast Fourier transform, and let the computer plot the Fourier components as a function of k. What one gets is $\tilde{g}(k)$, which is a completely different function from $g(x)$. However, it contains exactly the same information as $g(x)$, since it is merely the Fourier transform of $g(x)$ and its inverse Fourier transform is $g(x)$.

How about overtones? To answer, we must consider the frequency[6] of each mode, $\omega(k)$. Thus, in addition to $g(x)$ or $\tilde{g}(k)$, $h(\omega) \equiv \tilde{g}(k(\omega))$ is a good representation.

The "string vibration vector" is given a symbol $\tilde{\vec{g}}$ here: since the vector here is **wave as vector**, distinguishable from the vector of a one particle Newtonian mechanics, e.g. $\vec{J}$ in the first example. The squiggly vector notation is invented here for a temporary use, to emphasize the vector nature of wave. As we shall see shortly, the dimensionality of the vector space in the current example is infinite, not finite as would be for any vector involved in the description of a single particle Newtonian mechanics.

[6]In this course, if unspecified, frequency means *angular* frequency. This is a widely used convention, due to ω being a much more frequently used quantity than the linear frequency $\nu = \omega/(2\pi)$.

At this point, it is worth noting that the mathematics of wave in Newtonian mechanics is much like the mathematics of one particle in QM. The notable differences are (1) the equations of motions are different and (2) in Newtonian mechanics, one *does not have to* use complex numbers, while in QM, complex numbers appear at the equation of motion level, and thus cannot be avoided. Regarding point (2), however, note that even in Newtonian mechanics, problems involving simple harmonic oscillators and waves tend almost always to use complex numbers, as it is quite sensible to take advantage of the ease of math that occurs when the problem is extended to the complex plane (e.g. the exponential function is much nicer to deal with than sine or cosine functions).

Therefore, from now on, we consider all representations, $g(x)$, $\tilde{g}(k)$ etc. to be complex.

Now, what is the basis set that gave the representation $g(x)$ for $\tilde{g}$? The answer is a set of vectors $\{\tilde{x}_0\}$ that can be defined by the Dirac-orthonormality condition (cf. Eq. 1.10) alone:

$$\tilde{x_1}^\dagger \tilde{x_0} = \delta(x_1 - x_0) \qquad 0 \le x_0, x_1 \le L \tag{2.10}$$

At this point, let us switch to the Dirac notation, writing $|g\rangle$ instead of $\tilde{g}$. The fact is that the Dirac notation is applicable to *any* vector, and so it makes a perfect sense to use it here, or even for the previous example. Having said this, the convention is to use the Dirac notation only in QM, and so you are advised to be aware that using the Dirac notation for classical physics is uncommon.

In any case, **the Dirac notation makes things much nicer.** The above equation becomes (cf. Eq. 1.10)

$$\langle x_1 | x_0 \rangle = \delta(x_1 - x_0) \tag{2.11}$$

The x *representation* of $|g\rangle$ is defined by

$$g(x) \equiv \langle x | g \rangle \tag{2.12}$$

In particular, for $|x_0\rangle$, we get

$$x_0(x) \equiv \langle x | x_0 \rangle = \delta(x - x_0) \tag{2.13}$$

where the last step is the consequence of putting $x_1 = x$ in Eq. 2.11. Thus, we see that the ket $|x_0\rangle$ corresponds to the Dirac delta function at $x = x_0$. How many of

these basis kets exist? Infinitely many, since we can take x_0 to be any value from 0 to L. This is why the vector space considered here is infinite dimensional, as mentioned already in page 6.

In this case, the resolution of identity (Eq. 1.13) can be demonstrated explicity, in a trivial manner, establishing $\{|x\rangle\}$'s as a **natural basis** of this problem. Consider the operator[7]

$$\hat{R} \;=\; \int dx\, |x\rangle\langle x| \tag{2.14}$$

Our goal is to show that $\hat{R} = 1$. To do this, consider

$$\langle f|\,\hat{R}\,|g\rangle \;=\; \int dx\, \langle f\,|\,x\rangle\,\langle x\,|\,g\rangle \tag{2.15}$$

$$= \int dx\, f(x)^* g(x) \qquad \text{using Eq. 2.12 and its complex conjugate} \tag{2.16}$$

$$= \langle f\,|\,g\rangle \qquad \text{the inner product of two complex functions} \tag{2.17}$$

Thus, we see that $\langle f|\,\hat{R}\,|g\rangle = \langle f\,|\,g\rangle$ for *any* waves/functions f and g, and so we conclude $\hat{R} = 1$, proving **the resolution of identity**

$$1 \;=\; \int dx\, |x\rangle\langle x| \tag{2.18}$$

Using this identity, the representation $g(x)$ is recognized simply as the coefficient of the vector $|g\rangle$ along the $|x\rangle$ direction:

$$|g\rangle \;=\; 1\cdot|g\rangle \;=\; \int dx\, |x\rangle\langle x\,|\,g\rangle \;=\; \int dx\, g(x)\,|x\rangle \tag{2.19}$$

Now, let us adopt the Dirac notation *even for the previous example*. Then, Eqs. 2.8, 2.9 can be rewritten nicely as

$$|J\rangle \;=\; \sum_i |e_i\rangle\langle e_i\,|\,J\rangle \;=\; \sum_i J_i\,|e_i\rangle\,, \quad J_i \equiv \langle e_i\,|\,J\rangle$$

I hope you will agree with me that the Dirac notation makes it extremely clear how any vector can be resolved into components, i.e. how they can be represented. **This is one of the aspects that make the Dirac notation so elegant and powerful.**

[7] All integrals appearing in this lecture are definite integrals, with the range being the entire allowed range, even if the integration range is omitted (for brevity). By the same token, the integrals in LN 1 (e.g., in Eqs. 1.11, 1.13, 1.14) were also definite integrals, covering the entire range of e.

Change of basis is also easy to deal with, if we use the Dirac notation. Here, we will make the guitar string infinite in length, allowing the string to exist from $x = -\infty$ to ∞, to make the problem simple. Then, the k value allowed for the string is unrestricted, i.e. continuous from $-\infty$ to ∞.

$$\langle k \,|\, k' \rangle = \delta(k - k') \qquad \text{Orthonormality} \qquad (2.20)$$

$$1 = \int dk \, |k\rangle \langle k| \qquad \text{Resolution of identity} \qquad (2.21)$$

Here, $\langle x \,|\, k \rangle = A \exp(ikx)$ and the normalization constant A can be determined by the orthonormality as $A = 1/\sqrt{2\pi}$ using the well-known formula

$$\int_{-\infty}^{\infty} dx \, \exp\left[i(k - k')x\right] = 2\pi \, \delta(k - k') \qquad (2.22)$$

and so we get

$$\langle x \,|\, k \rangle = \frac{1}{\sqrt{2\pi}} \exp(ikx) \qquad (2.23)$$

We have just showed that if the k-space representation of a wave is a Dirac delta function ($\langle k' \,|\, k \rangle = \delta(k' - k)$ from Eq. 2.20), then its x-space representation is $\langle x \,|\, k \rangle = \frac{1}{\sqrt{2\pi}} \exp(ikx)$, providing the most basic example of the Heisenberg uncertainty principle.

Now, consider the k-space representation $\tilde{g}(k) = \langle k \,|\, g \rangle$ for an arbitrary wave $|g\rangle$. By inserting the resolution of identity for the x-basis, we get

$$\tilde{g}(k) = \langle k \,|\, g \rangle = \int_{-\infty}^{\infty} dx \, \langle k \,|\, x \rangle \langle x \,|\, g \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx \, \exp(-ikx) g(x) \qquad (2.24)$$

Similarly, $g(x)$ can be expressed in terms of $\tilde{g}(x)$ as

$$g(x) = \langle x \,|\, g \rangle = \int_{-\infty}^{\infty} dk \, \langle x \,|\, k \rangle \langle k \,|\, g \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk \, \exp(ikx) \tilde{g}(k) \qquad (2.25)$$

These two formulae are, of course, Fourier and inverse Fourier transformations. **How naturally they just *flow out to your attention* from the Dirac notation!**

Before we move onto the next example, we note that all numbered equations, Eqs. 2.11 through 2.25, are applicable to QM. We simply have to imbue them with the following re-interpretation. In QM, $|g\rangle$ would stand for a quantum mechanical state, and $|x\rangle$ can be recognized as eigenstate of the position operator and $|k\rangle$ as eigenstate of the momentum ($p = \hbar k$) operator.

Electron in a quantum corral – vector for a quantum particle

Having gone through the previous two examples, the QM case of electron in a nanometer-scale quantum corral made by Fe atoms is not so difficult to analyze from the vector point of view.

The first thing to note is that **the electron itself is a vector**, where "vector" in this case is analogous to the guitar string wave, rather than to the impact vector, of the two examples that we just finished discussing. Thus, the expression **"electron = vector"** is, of course, a re-phrasing of the good old **wave-particle duality**. It stands for nothing else than the fact an electron acts like it is a wave (as does any other quantum particle such as photon or neutron).

In this case, the vector $|\Psi\rangle$ is none other than the state of the electron. The *real space representation*

$$\langle \vec{x} \,|\, \Psi \rangle \;=\; \Psi(\vec{x}) \tag{2.26}$$

is commonly referred to as *the wave function*. The *wave vector space representation* (or the *momentum space representation*) is given by

$$\left\langle \vec{k} \,\middle|\, \Psi \right\rangle \;=\; \tilde{\Psi}(\vec{k}) \tag{2.27}$$

It is "the wave function in the $\vec{k}$ space." In the current example, the orthonormality and the resolution of identity for the real space basis is given by

$$\langle \vec{x} \,|\, \vec{x}' \rangle \;=\; \delta(\vec{x} - \vec{x}') \;=\; \delta(x - x')\delta(y - y') \qquad \vec{x} = x\hat{x} + y\hat{y},\ \vec{x}' = x'\hat{x} + y'\hat{y} \tag{2.28}$$

$$1 \;=\; \int d\vec{x}\, |\vec{x}\rangle \langle \vec{x}| \qquad d\vec{x} = dxdy \tag{2.29}$$

Note that the multi-dimensional Dirac delta function $\delta(\vec{x})$ or the multi-dimensional infinitesimal $d\vec{x}$ splits into products of components.

The basis states $\{|\vec{x}\rangle\}$ are eigenstates of the operator $\hat{\vec{x}}$. This is clear since

$$\hat{\vec{x}} \,|\vec{x}'\rangle \;=\; \vec{x}\,|\vec{x}'\rangle \;=\; \vec{x}'\,|\vec{x}'\rangle \tag{2.30}$$

where the last step arises since $|\vec{x}'\rangle \doteq \delta(\vec{x} - \vec{x}')$ in real space.

This is to be expected from the following property.

Eigenstates of *any* observable form a natural basis.

This is due to the fundamental hypothesis of QM that an observable corresponds to a Hermitian operator (page 7 of LN 1), and the property of a Hermitian operator (page 8 of LN 1). Also, note that, eigenstates = eigenvectors $\doteq$ eigenfunctions in QM, since state = vector $\doteq$ wave function.

Considering the (approximate) rotational symmetry, the above Cartesian decomposition is arguably not the best way. Instead, we can take $|r,\theta\rangle$ as the basis state. In this case, we have

$$\langle r,\theta\,|\,r',\theta'\rangle = \delta(r-r')\delta(\theta-\theta') \tag{2.31}$$

$$1 = \int_0^{2\pi} d\theta \int_0^{\infty} dr\ |r,\theta\rangle\langle r,\theta| \tag{2.32}$$

How about the $\vec{k}$-representation? Taking the eigenstates of $\hat{\vec{k}} = \hat{\vec{p}}/\hbar$, we get

$$\left\langle \vec{k}\,\middle|\,\vec{k}'\right\rangle = \delta(\vec{k},\vec{k}') = \delta(k_x-k_x')\delta(k_y-k_y') \tag{2.33}$$

$$1 = \int d\vec{k}\,\left|\vec{k}\right\rangle\left\langle\vec{k}\right| \qquad d\vec{k} = dk_x dk_y \tag{2.34}$$

Note that the $\vec{k}$-representation of $\left|\vec{k}\right\rangle$ is a Dirac delta function according to Eq. 2.33. How about the $\vec{x}$-representation of it? The eigenstate of $\hat{\vec{k}} = \hat{\vec{p}}/\hbar$ is a plane wave, since $\hat{\vec{p}} = -i\hbar\frac{\partial}{\partial\vec{x}}\left(\equiv -i\hbar\vec{\nabla}\right)$ *in the* $\vec{x}$*-representation*[8], $\left\langle\vec{x}\,\middle|\,\vec{k}\right\rangle = A\exp(i\vec{k}\cdot\vec{x})$, in the real space, and Eq. 2.33 determines the normalization constant $A = 1/(2\pi)$.

[8]More precisely speaking, $\left\langle\vec{x}\,\middle|\,\hat{\vec{p}}\,\middle|\,\vec{x}'\right\rangle = \delta(\vec{x}-\vec{x}')\left(-i\hbar\frac{\partial}{\partial\vec{x}}\right)$ (i.e., $\hat{\vec{p}}$ is a *diagonal matrix*, in the $\vec{x}$-representation, with diagonal elements $= -i\hbar\frac{\partial}{\partial\vec{x}}$), which means that $\left\langle\vec{x}\,\middle|\,\hat{\vec{p}}\,\middle|\,\Psi\right\rangle = \int d\vec{x}'\left\langle\vec{x}\,\middle|\,\hat{\vec{p}}\,\middle|\,\vec{x}'\right\rangle\langle\vec{x}'\,|\,\Psi\rangle = -i\hbar\frac{\partial\Psi(\vec{x},t)}{\partial\vec{x}}$.

Representation

Consider a natural basis set $\{|e\rangle\}$, which could be a discrete set, a continuous set, or a mixed set (Eqs. 1.12-1.14). The representation of a vector $|\Psi\rangle$ is given by $\langle e|\Psi\rangle$. The representation of an operator $\hat{O}$ is given by $\langle e|\hat{O}|e'\rangle$.

Representation = Components
Representation of a state vector = Wave function

Here, e can be any quantum numbers associated with an observable: spatial coordinate $\vec{x}$ (for which the observable is position), wave vector $\vec{k}$ (for which the observable is momentum/$\hbar$), energy E, or angular momentum quantum numbers $e = (l, m)$, etc. The $\vec{x}$ representation is the $\vec{x}$ space wave function, the most commonly used wave function by far. The $\vec{k}$ representation is called the $\vec{k}$ space wave function.function.

There is no $\vec{x}$ or $\vec{k}$, etc., dependence in $|\Psi\rangle = |\Psi(t)\rangle$. Such dependence arises only when represented:
$\langle \vec{x}|\Psi\rangle = \Psi(\vec{x}, t)$, $\left\langle \vec{k}\middle|\Psi\right\rangle = \Psi(\vec{k}, t)$.

This is as obvious as saying that, in example 1, $\vec{J}$ does not depend on i, the component index. A state $|\Psi\rangle$ is just that, a state *as a whole*: the *component index* such as $\vec{x}$ or $\vec{k}$ is not applicable for $|\Psi\rangle$. Of course, $\Psi(\vec{x})$ is the most commonly used representation. However, it is important to know how to switch between representations.

Assuming a radial potential well, the angular momentum representation is interesting as well. For an electron in two dimensions, such as this problem, the angular momentum $\hat{L}_z$ is the only angular momentum to consider, where z direction is defined to be perpendicular to the plane. Since $\hat{L}_z = -i\hbar\frac{\partial}{\partial\theta}$, in the (r, θ) representation[9], we have

[9]Similarly as for $\hat{\vec{p}}$, the more precise statement is that $\langle r, \theta|\hat{L}_z|r', \theta'\rangle = \delta(r - r')\delta(\theta - \theta')\left(-i\hbar\frac{\partial}{\partial\theta}\right)$.

$$\langle m \,|\, m' \rangle \;=\; \delta_{m,m'} \qquad\qquad m, m' \in \mathbb{Z} \text{ (integers)} \tag{2.35}$$

$$1 \;=\; \sum_{m \in \mathbb{Z}} |m\rangle \langle m| \tag{2.36}$$

Here, $|m\rangle$ is defined as an eigenstate of $\hat{L}_z$:

$$\hat{L}_z \,|m\rangle \;=\; m\hbar \,|m\rangle \tag{2.37}$$

In the θ representation, we get

$$-i\hbar \frac{\partial}{\partial \theta} \langle \theta \,|\, m \rangle \;=\; m\hbar \,\langle \theta \,|\, m \rangle \tag{2.38}$$

which shows that the wave function $\langle \theta \,|\, m \rangle$ is given by

$$\langle \theta \,|\, m \rangle \;=\; \frac{1}{\sqrt{2\pi}} \exp(im\theta) \tag{2.39}$$

Here, the normalization factor $1/\sqrt{2\pi}$ ensures that $\langle m \,|\, m \rangle = \int_0^{2\pi} d\theta \, |\langle \theta \,|\, m \rangle|^2 = 1$.

If the confining potential by the corral is given, then the energy eigenvalues of the energy can be solved for and the (Dirac-)orthonormality and the resolution identity can be written in terms of them. For this problem, the energy eigenstates will likely be of mixed character: some bound states at low energy and continuum of states above a threshold.

As we end this example, let us note that this example was for particle in two dimensions. The extension to other spatial dimensions should be straightforward. In particular, the case of the simpler one dimensional case has been already presented in the previous example, whose **Eqs. 2.11 through 2.25 are applicable to QM without modification)**.

Another notable thing in the context of representation is the following.

Inner product

The inner product between two vectors $\langle \Psi_1 | \Psi_2 \rangle$ is invariant upon change of basis, for any vectors $|\Psi_1\rangle$ and $|\Psi_2\rangle$, assuming both basis sets involved are natural basis sets (LN 1, page 8). So, for instance, in a one dimensional problem where x is unrestricted (and so k is continuous) and the energy spectrum is discrete,

$$\langle \Psi_1 | \Psi_2 \rangle = \int_{-\infty}^{\infty} dx\, \Psi_1^*(x)\Psi_2(x) \qquad \Psi(x) \equiv \langle x | \Psi \rangle \qquad (2.40)$$

$$= \int_{-\infty}^{\infty} dk\, \Psi_1^*(k)\Psi_2(k) \qquad \Psi(k) \equiv \langle k | \Psi \rangle \qquad (2.41)$$

$$= \sum_n \Psi_1^*(n)\Psi_2(n) \qquad \Psi(n) \equiv \langle n | \Psi \rangle,\ \hat{H}|n\rangle = E_n |n\rangle \qquad (2.42)$$

The last equation can be rewritten as $\sum_n C_{1,n}^* C_{2,n}$, following the common notation $|\Psi\rangle = \sum_n C_n |n\rangle$, where $C_n = \langle n | \Psi \rangle$. That is, the coefficient C_n is in fact *the energy space wave function*.

More generally, any scalar quantity is invariant under change of natural basis.

For example, $\langle \Psi_1 | \hat{O} | \Psi_2 \rangle$ is such a scalar quantity.

2.1.2 Resolution of identity

In the above, we saw that the (Dirac-)orthonormality and the resolution of identify are basic starting points in the theory of representation.

The resolution of identity expresses the fact that the orthonormal basis is **complete**, i.e. it spans the entire Hilbert space. The resolution of identity can be proved in purely mathematical terms in some cases. However, the physical basis for it must be appreciated, first and foremost.

In QM, we take a basis from eigenstates of an observable. Then the resolution of identity, is the consequence of the following reasonable assumptions: when an observable is measured on an arbitrary state, then (1) it must produce *some* result and (2) the total probability of all possible outcomes is always unity.

Each individual term that enters the resolution of identity, $|e\rangle\langle e|$, is called projection operator:

$$\hat{P}_e \equiv |e\rangle\langle e| \qquad \textbf{projection operator} \tag{2.43}$$

The expectation value for the projection operator for a given state is given by

$$\left\langle \Psi \left| \hat{P}_e \right| \Psi \right\rangle = |\langle e | \Psi \rangle|^2 \tag{2.44}$$

which, according to the measurement theory of QM (LN 1, pages 9,10), gives the probability if e is discrete, or the probability density if e is continuous, for the measurement of $\hat{e}$ to give the value e.

2.1.3 Schrödinger equation, revisited

As promised in the last LN, we will now "derive" Eq. 1.27. Please re-read contents of the box in page 12 if you have not done so, as they are essential for this discussion. We start from the Schrödinger equation, Eq. 1.15,

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

Now, we multiply this equation by $\langle \vec{x}|$ from the left.

$$i\hbar \left\langle \vec{x} \left| \frac{d}{dt} \right| \Psi(t) \right\rangle = \left\langle \vec{x} \left| \hat{H} \right| \Psi(t) \right\rangle \tag{2.45}$$

Taking $\vec{x}, t$ as the set of independent variable of calculus, we see that d/dt can be taken out of the bra-ket, *if* we change it to the partial time derivative.

$$i\hbar \frac{\partial}{\partial t} \langle \vec{x} | \Psi(t) \rangle = i\hbar \frac{\partial \Psi(\vec{x}, t)}{\partial t} = \left\langle \vec{x} \left| \hat{H} \right| \Psi(t) \right\rangle \tag{2.46}$$

The right hand side can be expanded as (by inserting the resolution of identity)

$$\left\langle \vec{x} \left| \hat{H} \right| \Psi(t) \right\rangle = \int d\vec{x}' \left\langle \vec{x} \left| \hat{H} \right| \vec{x}' \right\rangle \langle \vec{x}' | \Psi(t) \rangle$$

We shall assume that $\hat{H}$ is diagonal in the real space representation, as would be normally the case (cf. footnote 8). Then,

$$\left\langle \vec{x} \left| \hat{H} \right| \vec{x}' \right\rangle = \delta(\vec{x} - \vec{x}')\, H(\vec{x}, \vec{p}, t)|_{\vec{p} = -i\hbar \frac{\partial}{\partial \vec{x}}}$$

where H on the right hand side is the classical Hamiltonian function, with $\vec{p}$ replaced by $-i\hbar\frac{\partial}{\partial\vec{x}}$. Plugging in this result to the equation before it, and effecting the integral over $\vec{x}'$, we get

$$\left\langle \vec{x} \left| \hat{H} \right| \Psi(t) \right\rangle = H(\vec{x},\vec{p},t)|_{\vec{p}=-i\hbar\frac{\partial}{\partial\vec{x}}} \Psi(\vec{x},t)$$

Thus, Eq. 2.46 becomes

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

which is Eq. 1.27. While this equation may be more familiar, Eq. 1.15 is more general.

Finally, note that in discussing the Schrödinger equation here, we have tacitly assumed that there is no other dynamical variable than $\vec{x}$, $\vec{p}$. This would not be true if the spin dynamics is involved. In that case, $H = H(\vec{x},\vec{p},\vec{S},t)$, where $\vec{S}$ is the spin vector, each component of which is a spin matrix (like a Pauli matrix).

2.1.4 Some important unitary operators

We have already identified

$$\hat{\mathcal{U}}(dt) = 1 - i\frac{\hat{H}}{\hbar}dt \tag{1.18}$$

as the infinitesimal **time evolution operator** by $t \to t + dt$.

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

For a finite translation in time Δt, we get

$$\hat{\mathcal{U}}(\Delta t) = \exp\left(-i\frac{\hat{H}\Delta t}{\hbar}\right) \quad \textit{only if } \hat{H}(\hat{x},\hat{p},t) = \hat{H}(\hat{x},\hat{p}), \text{ i.e. no explicit } t \text{ in } \hat{H} \tag{1.22}$$

Let us ask a question: can we find a similar operator for translation or rotation? The answers are

$$\hat{\mathcal{T}}(d\vec{x}) = 1 - i\frac{\hat{\vec{p}}}{\hbar}\cdot d\vec{x} \qquad \text{infinitesimal } \textbf{translation operator} \tag{2.47}$$

$$\hat{\mathcal{R}}(d\theta) = 1 - i\frac{\hat{L}_\theta}{\hbar}d\theta \qquad \text{infinitesimal } \textbf{rotation operator} \tag{2.48}$$

Here, $\hat{\vec{p}}$ and $\hat{L}_\theta$ are momentum and angular momentum operators, respectively. Unlike $\hat{H}$, they can never contain any explicit t dependence (by definition), and so the finite

translation operator and the finite rotation operator *are always* given by (doing just like for the derivation of Eq. 1.22)

$$\hat{\mathcal{T}}(\Delta\vec{x}) \;=\; \exp\left(-i\frac{\hat{\vec{p}}}{\hbar}\cdot\Delta\vec{x}\right) \qquad \textbf{translation operator} \qquad (2.49)$$

$$\hat{\mathcal{R}}(\Delta\theta) \;=\; \exp\left(-i\frac{\hat{L}_\theta}{\hbar}\Delta\theta\right) \qquad \textbf{rotation operator} \qquad (2.50)$$

The rotation operator is applicable even for the spin rotation. In that case, $\hat{L}_\theta$ must be identified as $\hat{S}_\theta$, the spin angular momentum associated with the θ rotation using the right hand/screw rule.

It is easy to prove the forms of the above translation operator and rotation operator. It suffices to prove the infinitesimal form. Here, we will show how $\hat{\mathcal{R}}(d\theta)$ can be derived, and the derivation for $\hat{\mathcal{T}}(d\theta)$ is completely analogous.

Given a wave function $\Psi(\theta)$ (other coordinates, if any, are suppressed, as they are irrelevant for this discussion), we like to find the operator $\hat{\mathcal{R}}$ such that

$$\left\langle \theta \left| \hat{\mathcal{R}}(d\theta) \right| \Psi \right\rangle \;=\; \langle\, \theta - d\theta \,|\, \Psi \rangle \;=\; \Psi(\theta - d\theta) \qquad (2.51)$$

Now, the left hand side $= \left\langle \theta \left| \hat{\mathcal{R}}(d\theta) \right| \Psi \right\rangle = \mathcal{R}(d\theta)\Psi(\theta)$ where $\mathcal{R}(d\theta)$ is the diagonal element of the θ representation of $\hat{\mathcal{R}}$. On the other hand, the right hand side $= \Psi(\theta) - d\theta\frac{\partial\Psi}{\partial\theta}$. Equating these two, we get

$$\mathcal{R}(d\theta) \;=\; 1 - d\theta\,\frac{\partial}{\partial\theta}$$

From page 12, we see that $\frac{\partial}{\partial\theta}$ is the representation of $i\frac{\hat{L}_\theta}{\hbar}$ (note that $\hat{L}_z$ there is the same as $\hat{L}_\theta$ here). Thus, we get an operator equality[10], $\hat{\mathcal{R}}(d\theta) = 1 - id\theta\,\hat{L}_\theta/\hbar$, proving Eq. 2.48.

Our findings here are summarized as follows. The Hamiltonian is the generator of the time evolution. The momentum is the generator of the translation. The angular momentum is the generator of the rotation.

These facts have important implications on conservation principles, as we shall see shortly.

[10]To prove that two operators are identical, it is sufficient to prove that they are identical in one representation.

2.1.5 Measured quantities

In QM, the quantities that can be compared with experiments have the form of, or are derived from quantities of the form of,

$$\left\langle \Psi_1 \left| \hat{O} \right| \Psi_2 \right\rangle \tag{2.52}$$

Quantities of this form are generally referred to as **matrix elements**, for the reason that an operator $\hat{O}$ is a square matrix, and since the above expression correspond to a component/element of the $\hat{O}$ matrix.

2.1.6 Heisenberg equation of motion

This is a topic for your optional reading. I will not mention this topic at all, outside this section! This topic is not difficult and it is rather nice to learn, but we do not really need it for this course. Knowing the Heisenberg equation of motion will enrich your QM knowledge.

In the **Schrödinger picture**, we consider states as evolving in time, while operators such as $\hat{x}$, $\hat{p}$, etc. fixed in time. Thus, the above matrix element expression can be written as (ignoring spin, for brevity)

$$\left\langle \Psi_1(t) \left| \hat{O}(\hat{x}, \hat{p}, t) \right| \Psi_2(t) \right\rangle \tag{2.53}$$

Here, the t dependence on $\hat{O}$ is for any extrinsic t dependence, e.g., when the Hamiltonian is driven by an external agent, that might exist. For a closed system, there would be no t dependence for operators.

So, in the Schrödinger picture, we consider all basic observable variables[11] (position, momentum, spin angular momentum, spin angle, etc.) time independent, while states are considered time dependent.

In the **Heisenberg picture**, the opposite is true. Operators are considered time-dependent, while states are considered time-independent.

Note that in terms of the time evolution operator $\hat{\mathcal{U}}(t)$, the above matrix element can be written as

$$\left\langle \hat{\mathcal{U}}(t)\Psi_1(t=0) \left| \hat{O} \right| \hat{\mathcal{U}}(t)\Psi_2(t=0) \right\rangle \;=\; \left\langle \Psi_1(t=0) \left| \hat{\mathcal{U}}(t)^\dagger \hat{O} \hat{\mathcal{U}}(t) \right| \Psi_2(t=0) \right\rangle \tag{2.54}$$

[11] By "basic observables," what I mean are those dynamical variables that define the "phase space" in classical mechanics.

From this, the operator in the Heisenberg picture $\hat{O}_h$ is defined as

$$\hat{O}_h \equiv \hat{\mathcal{U}}(t)^\dagger \hat{O} \hat{\mathcal{U}}(t) \tag{2.55}$$

while a state is defined, in the Heisenberg picture, as

$$|\Psi\rangle_h \equiv |\Psi(t=0)\rangle \tag{2.56}$$

Therefore, the measured quantity, i.e. the matrix element, is invariant when we go from the Schrödinger picture to the Heisenberg picture. That is, ${}_h\langle\Psi_1|\hat{O}_h|\Psi_2\rangle_h$ is exactly the same as the above matrix element (Eq. 2.53), by construction.

The **Heisenberg equation of motion** can be derived by considering the infinitesimal time evolution operator $\hat{\mathcal{U}}(dt)$ (Eq. 1.18) in Eq. 2.55. For complete derivation, see `http://griffin.ucsc.edu/teaching/08Q2-139A/download/L8%20-%20Formalism.pdf`. The result is

$$\frac{d\hat{O}_h}{dt} = \frac{i}{\hbar}\left[\hat{H}_h, \hat{O}_h\right] + \left(\frac{\partial \hat{O}}{\partial t}\right)_h \qquad \text{Note that } \left[\hat{H}_h, \hat{O}_h\right] = \left[\hat{H}, \hat{O}\right]_h. \tag{2.57}$$

Plugging in $\hat{H}$ for $\hat{O}$, we see that $\hat{H}_h$ is time-invariant if $\hat{H}$ does not have any explicit time dependence. This is precisely analogous to the result in classical mechanics: the homogeneity of time or the time invariance leads to the conservation of energy/Hamiltonian. For other dynamical variable that does not have any explicit time dependence, it is invariant if it commutes with $\hat{H}$. These results are familiar ones from QM I, and are re-summarized in the next section from the Schrödinger picture point of view.

Within the Heisenberg picture, the role of state is merely specifying the initial condition, and no more question is asked about the state. Questions that are asked and answered in this picture are "how does the momentum/position/energy/any-observable change over time?". These questions are apparently very similar to those asked in Newtonian mechanics. Of course, the difference here is the devilish detail that the observable is a matrix, a very large matrix as a rule, in QM.

2.1.7 Symmetry and conservation

An observable $\hat{O}$ is said to be conserved if

$$\frac{d}{dt}\left\langle\Psi\left|\hat{O}\right|\Psi\right\rangle = 0 \qquad \text{for } \textit{any} \text{ state } |\Psi\rangle \tag{2.58}$$

The expression $\langle \Psi | \hat{O} | \Psi \rangle$ defines the **expectation value** of $\hat{O}$ for $|\Psi\rangle$. Using the infinitesimal time evolution operator (Eq. 1.18), one can see that

$$d\langle \Psi | \hat{O} | \Psi \rangle = \langle \Psi | \hat{\mathcal{U}}^\dagger(dt)\, \hat{O}(t+dt)\, \hat{\mathcal{U}}(dt) | \Psi \rangle - \langle \Psi | \hat{O}(t) | \Psi \rangle \tag{2.59}$$

where on the right hand side, $\hat{O}(t)$ is any explicit dependence of $\hat{O}$ on t, while the dependence of $\hat{O}$ on basic observables such as $\hat{x}$ and $\hat{p}$ are not written down explicitly, to save space. Using Eq. 1.18, one can derive the following principles.

Principles of conservation

Energy conservation

If $\hat{H}$ does not explicitly depend on t, then $\hat{H}$ is conserved.

Conservation in general

If $\hat{O}$ does not explicitly depend on t and if $\left[\hat{O}, \hat{H}\right] = 0$, then $\hat{O}$ is conserved.

Let us assume that the first condition is satisfied. Then, $\hat{H}$ is conserved, and its eigenstates are **stationary states**, since each eigenstate simply acquires a phase factor $\exp(-iE_n t/\hbar)$ (LN 1, page 11).

In addition, suppose that there is a **conserved observable $\hat{O}$.** This means that $\left[\hat{H}, \hat{O}\right] = 0$, and thus, **an energy eigenstate can be taken to be a simultaneous eigenstate of $\hat{O}$** (see LN 1, page 9, **Compatible operators**).

Now, as an example, suppose that $\hat{L}_\theta$ is conserved, i.e. we find that $\left[\hat{H}, \hat{L}_\theta\right] = 0$, where $\hat{L}_\theta$ is the angular momentum operator discussed in Section 2.1.4. Since the rotational operator $\hat{\mathcal{R}}(\theta)$ is a function of $\hat{L}_\theta$ (Eq. 2.48 or 2.50), it follows that the similarity transformation $\hat{\mathcal{R}}\hat{H}\hat{\mathcal{R}}^\dagger = \hat{\mathcal{R}}\hat{\mathcal{R}}^\dagger\hat{H} = \hat{H}$, where in the last step the unitary nature of the rotation operator is used. So, we see that the condition $\left[\hat{H}, \hat{L}_\theta\right] = 0$ is equivalent to the **rotational invariance of $\hat{H}$, i.e. the rotational invariance of the system. So, just like in classical mechanics, a rotationally invariant system conserves the angular momentum. Similarly, a translationally invariant system conserves the linear momentum.**

Other symmetries quite useful to consider include the reflection symmetry ("parity") and the inversion symmetry.

In approaching a QM problem, the first thing to do is to notice valid symmetries. Then, energy eigenstates can be written down as simultaneous eigenstates of all valid and mutually compatible symmetries. For instance, if there is a parity symmetry, then each eigenstate can be written down as an even function or an odd function of the corresponding reflection axis. Or, if the system has a rotational symmetry, then eigenstates can be written down as angular momentum eigenstates (spherical harmonics). As a rule, these steps greatly simplify the task of solving the Hamiltonian eigenvalue problem, which is the first thing to do for any QM problem.

2.1.8 Uncertainty principle

For any given two observable $\hat{A}$ and $\hat{B}$, and any physical state[12],

$$\Delta A \Delta B \geq \left| \frac{\langle [\hat{A}, \hat{B}] \rangle}{2i} \right| \tag{2.60}$$

where on the right hand side the quantity inside the absolute value sign is a real number, $\Delta O \equiv \sqrt{\langle \hat{O}^2 \rangle - \langle \hat{O} \rangle^2}$ is the uncertainty of $\hat{O}$, and $\langle \hat{O} \rangle$ is the short hand for the expectation value $\langle \Psi | \hat{O} | \Psi \rangle$.

If the two variables are canonical conjugate to each other, then we have $[\hat{A}, \hat{B}] = \pm i\hbar$, and so we get

$$\Delta A \Delta B \geq \frac{\hbar}{2} \tag{2.61}$$

The qualitative meaning of the uncertainty principle is that **if two observables $\hat{A}$ and $\hat{B}$ are incompatible (i.e. have a non-zero commutator), then the wave function in the A representation and the wave function in the B representation cannot be made very sharp simultaneously.**

While the energy and the time are not conjugate observables in QM (because t is not an observable; the energy and the time are still "conjugate"), this qualitative statement is still valid for them. If the wave function is localized in energy, then the wave function is extended in time – the state has a long life time during which it does not change appreciably. If the wave function is extended in energy, then it will

[12] Recall that a physical state is a normalizable state (LN 1, page 4).

quickly change in time – the state will decay very rapidly. With a properly defined Δt, one can show $\Delta E \Delta t \geq \hbar/2$ (http://griffin.ucsc.edu/teaching/08Q2-139A/download/L8%20-%20Formalism.pdf).