Quantum Mechanics

Elad Harel

LECTURE 1
Review of Quantum Mechanics - Part I

1.1. Outline of Lecture

- State Space
- Operators and Observables
- Matrix Representation

With the exception of the classical mechanics discussion and a few examples and exercises, everything in this lecture is taken in one form or another from Sakurai's book, *Modern Quantum Mechanics*, which I highly recommend for any serious student of introductory quantum mechanics. However, I'm presenting things here out of order and filtered with respect to what I believe is important for understanding optical spectroscopy. I've included some derivations, but not all, and the interested reader should try them out on his or her own before looking them up. Derivations that I feel are too long or complex will be included in the supplementary notes for those interested. If you really want to understand this subject I highly urge you to do the exercises. Many of the exercises can be done with paper and pencil, although a few require Matlab. This is a good opportunity to learn Matlab if you don't know it already.

1.2. State Space

Quantum mechanics provides a framework in which we can describe the state of a system as we would describe a vector in a vector space. To make this leap from physical reality to abstract mathematics it is necessary to introduce a set of postulates which, through extensive experience, we come to believe as true. The role of quantum mechanics is not to explain *why* these postulates hold, only *how* they apply to the physical world. As our experiments

become more refined the postulates may change, but keep in mind that quantum mechanics is, arguably, the most validated scientific theory of all time. We won't cover the historical progression in which certain classical descriptions of physical realty were successfully challenged and explained (photoelectric effect, Compton effect, blackbox radiation, etc...) as there are numerous books on the subject.

This section will first describe the mathematical notation and tools necessary to deal with quantum mechanical systems. It will then cover the concept of measurement before venturing into a discussion on dynamics. Because the idea of angular momentum is central to an understanding of NMR and two-level systems (TLS), in general, we will pay special attention to this most important topic in Lecture 2.

1.2.1. Hilbert Space

A Hilbert space, $\mathbb{H}$, is a vector space with a complex scalar product which satisfies the property

$$\mu \cdot \upsilon = (\upsilon \cdot \mu)^* \qquad \mu, \upsilon \in \mathbb{H}$$

where * implies complex conjugate.

Example 1. Just as a refresher, the complex conjugate is constructed simply by transposing and conjugating (turning i to $-i$) and can be applied to any type of 'vector', e.g. vector, matrix, function (the word 'vector' is loosely defined as any object belonging to a vector space and, hence, does not necessarily imply the traditional 1xn vector we're used to thinking about). For example,

$$(1, 2+i, -3i)^* = \begin{pmatrix} 1 \\ 2-i \\ 3i \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1-i \\ 2+i & 0 \end{pmatrix}^\dagger = \begin{pmatrix} 1 & 2-i \\ 1+i & 0 \end{pmatrix}$$

$$\exp(ix^2)^* = \exp(-ix^2)$$

A notational note: For matrices we don't use the symbol *. Instead we use what's called the Hermitian adjoint which is labeled as †. The reason for this distinction will become more clear a bit later. $\diamondsuit$

Recall that a vector space is a linear space, meaning that any linear combination of vectors belonging to that vector space also belongs to that vector space. Mathematically, we write $a\mu + b\nu \in \mathbb{H}$ if $\mu, \nu \in \mathbb{H}$ where a, b are any complex numbers.

Furthermore, the scalar product of a vector with itself, which we call the norm, is always positive definite:

$$\|\mu\|^2 \equiv \mu \cdot \mu \geq 0, \qquad \mu \in \mathbb{H}$$

with $\|\mu\| = 0$ if and only if $\mu = 0$.

An example of a Hilbert space is the set of all square integrable functions, f, on the interval, $[a, b]$, along with the scalar product (operator)

$$f \cdot g = \int_a^b f(\mathbf{r}) g(\mathbf{r}) d\mathbf{r}$$

By square integrable, we mean a function where $f^2 = f \cdot f$ is normalizable (i.e. does not diverge).

Another more familiar example is simply the set of all vectors in three-dimensions (i.e. $\Re^3$) with the scalar product being defined as simply the dot product. Remember, the scalar product and vector space can be defined in many ways as long as the properties above are satisfied. The vector space can consist of vectors, matrices, function, etc... The scalar product can be any operation which results in a scalar as opposed to another "vector" and in which the above properties hold. Examples include the dot product, integral, etc... The space we call a Hilbert space.

Exercise 1. Why is the trace not a suitable scalar product in the vector space consisting of all 2x2 matrices?

$$A \cdot B = tr(AB)$$

♣

It is also important to note that the Hilbert space may be finite dimensional or continuous. This distinction is crucial and determines what kind of representation is most suitable for the quantum system in question. We'll discuss this a bit later along with some concrete examples.

1.2.2. Dirac Notation

So what does the Hilbert space have anything to do with quantum mechanics and with physical reality? Quantum mechanics postulates that a physical state is represented by a state vector belonging to a Hilbert space. An example of a physical state may be an atom, a molecule, or a particular configuration of that atom or molecule. The atom may be in the ground state or in an electronically excited state. We'll discuss later what this actually means, but for now it suffices to say that the constituent parts of the atom have many degrees of freedom, and each configuration is what we call a physical state. For a molecule, as you can imagine, there are even more degrees of freedom and even more possible configurations.

One of the most important degrees of freedom that concerns us is spin angular momentum which we'll discuss a bit later.

This state vector is postulated to contain all the information about the physical state. To represent this "vector", Dirac introduced the bra and ket notation. These are just abstract mathematical constructs that represent the physical state of the system and they obey all the properties of vectors in a Hilbert space. So again, any linear combination of kets is itself a ket:

$$a|\alpha\rangle + b|\beta\rangle = |\gamma\rangle$$

Quantum mechanics postulates that only the 'direction' of the ket is of importance. That is, $c|\alpha\rangle$ is the same state as $|\alpha\rangle$. The bra is simply the complex conjugate transpose of the ket

$$\langle\alpha| = (|\alpha\rangle)^*$$

So now we have kets which we postulated contain complete information about the system and anything we could possible want to know about it. The set of all kets along with the scalar product define a Hilbert space, which, from now on, is called the ket space. The scalar product is written as follows

$$\langle\alpha|\beta\rangle$$

We have not yet provided a prescription for evaluating the inner product, but we do postulate two fundamental properties of it:

$$\langle\beta|\alpha\rangle = (\langle\alpha|\beta\rangle)^*$$

and

$$\langle\alpha|\alpha\rangle \geq 0$$

These are of course the properties we require of any scalar product, and are a direct consequence of our original postulate stating that the state vector belongs to a Hilbert space. It's often the case that you'll see one Quantum mechanical postulate dressed like another. In fact, there are very few of them just like in any other field (e.g. geometry) because through time they are reduced to only the most fundamental, which have no proof, but which all other knowledge can be built. Just for perspective, in Euclidean geometry there is a postulate that any two points make up a line. This statement can't be proved even though it appears completely trivial. But with this and the remaining postulates all of geometry can be constructed. The same with quantum mechanics!

Two additional definitions that are of importance are the notion of orthogonality and of the norm. Two kets are said to be orthogonal if

$$\langle\alpha|\beta\rangle = 0$$

The norm is given by

$$\|\alpha\|^2 = \langle\alpha|\alpha\rangle$$

Because only the direction of the ket is important we can conveniently normalize each ket by dividing by its norm.

1.2.3. Base kets

For all the remaining discussion we will assume that the Hilbert space is finite dimensional. Examples of infinite dimensional Hilbert spaces are those involving position and momentum states. We'll always be dealing with systems whose degrees of freedom can only take on discrete values and so our Hilbert space is finite dimensional. Again, this will become more clear once we discuss spin.

In any finite dimensional Hilbert space it is possible to find a set of vectors (states, kets, etc...) which form a complete set or basis. This means that any vector belonging to that Hilbert space can be written as a linear combination of these basis vectors, and, furthermore, that these basis vectors are orthonormal (i.e. orthogonal and normalized). Mathematically, we write

$$|\alpha\rangle = \sum_{a'} c_{a'}|a'\rangle$$

where $\{|a'\rangle\}$ is the set of basis kets and $c_{a'}$ are complex numbers. Because of the orthonormality of these kets we have

$$c_{a'} = \langle a'|\alpha\rangle$$

Exercise 2. Prove this. ♣

This gives

$$|\alpha\rangle = \sum_{a'} |a'\rangle\langle a'|\alpha\rangle$$

It then follows that

$$\sum_{a'} |a'\rangle\langle a'| = 1$$

where 1 is the identity operator (i.e. it acts onto a state vector, $|\alpha\rangle$, to give back $|\alpha\rangle$). This is known as the completeness relation and is extremely useful in quantum mechanics.

Exercise 3. Prove the following using the completeness relation:

$$\langle\alpha|\alpha\rangle = \sum_{a'} |c_{a'}|^2$$

Note the special case when $|\alpha\rangle$ is normalized. ♣

The usefulness of the basis kets is that once these are constructed, the expansion coefficient uniquely describe the state ket. In other words the basis kets provide a representation that allows us to deal with arbitrary states in a useful way. This will become more clear later on when we discuss matrix representation and operators.

1.3. Operators and Observables

We have so far constructed a very abstract formalism consisting of kets belonging to a Hilbert space. Now, we need a way to tie this into the real world in which we actually observe the properties of a system. In quantum mechanics the observation of a system is intrinsically tied to the state of the system. We are not simply observers, our action of observing changes the very state we're trying to observe, so quantum mechanics postulates. There is no such thing as the existence of a state independent of the act of observing it. I won't get into too philosophical a discussion of what this means and how to interpret it, but suffice it to say that observing a state necessarily disturbs it and so reality is only that of disturbed states. Weird, I know.

An operator, A, acts on a ket to produce another ket much like a matrix acts on a vector to produce another vector. Every observable, whether it be position, momentum, spin, vibrations, energy, etc... is represented by an operator (although for now we'll only deal with observables giving discrete values). An important set of kets associated with each operator are called eigenkets. As for matrices, whose eigenvectors when acted on the left by the matrix gives back the eigenvector with a proportionality factor called the eigenvalue, operators have eigenkets and corresponding eigenvalues:

$$A|a'\rangle = a'|a'\rangle$$

For every observable the set of eigenkets forms an orthonormal basis. This can easily be proved when we consider the fact that the Hilbert space is finite dimensional and therefore, the set of eigenkets, $\{|a'\rangle\}$, is finite. It is important to require that the observable be Hermitian which implies that

$$A^\dagger = A$$

where the $\dagger$ is the complex conjugate. The reason will become a little more clear later on when we relate the eigenvalues to physical observables.

Example 2. Later on we'll see that operators can be represented by matrices. The following is an example of a Hermitian matrix:

$$A = \begin{pmatrix} 3 & 1+i \\ 1-i & 5 \end{pmatrix}$$

Exercise 4. Find the eigenvalues and eigenvectors of A by hand and using Matlab. ♣

Exercise 5. Prove that the eigenvalues of any Hermitian operator are real. ♣

Theorem 1.1. *The eigenkets of any Hermitian operator corresponding to different eigenvalues are orthogonal.*

Proof. Let $|a'\rangle$ and $|a''\rangle$ be eigenkets of the Hermitian operator A corresponding to the eigenvalues a' and a'', respectively, where $a' \neq a''$. We have

$$A|a'\rangle = a'|a'\rangle$$
$$A|a''\rangle = a''|a''\rangle$$

Because A is Hermitian we must also have

$$\langle a'|A = a'\langle a'|$$

Recall that the eigenvalues of A are real which implies $a' = (a')^*$.
Therefore, we have

$$\langle a'|A|a''\rangle = a'\langle a'|a''\rangle$$

and

$$\langle a'|A|a''\rangle = a''\langle a'|a''\rangle$$

Since $a' \neq a''$, we must have

$$\langle a'|a''\rangle = 0$$

□

Lets summarize briefly what we've done and why we've done it. We introduced the notion of kets which we postulated contained all the information about our physical system. These kets we postulated belonged to a Hilbert space. In the physical world we must observe properties of these kets which we do by acting on them with operators that we postulate to be Hermitian. For each operator we can construct a set of basis kets from the eigenkets, and we know that any arbitrary ket can be written as a linear combination of these. What's missing is to somehow connect the observables of a system with the eigenkets. For that we make the following postulate which I remind you cannot be proven:

"A measuremnt always causes the system to jump into an eigenstate of the dynamical variable that is being measured" (Dirac 1958)

This is astounding. It means that for any observable we can construct the corresponding eigenkets and eigenvalues. The system which is typically in some arbitrary ket of the Hilbert space can be written as a linear combination of these eigenkets. When a measurement is made it causes the ket to 'collapse' into on of the eigenkets. The observable is then simply the eigenvalue of that eigenket. For every new measurement we construct a new observable (i.e. operator), find its eigenkets and eigenvalues and then write our state ket in terms of those. Now we understand why the observable must be Hermitian. It wouldn't make any sense for our observable (i.e. the eigenvalue) to be imaginary. It still doesn't answer the

question of how the system decides WHICH eigenstate to jump into. For this we'll need to discuss expectation values and ensembles.

Remark. For the sake of brevity I won't discuss why this fundamental postulate of quantum mechanics makes sense. Its a great leap of logic and one that's been tested time and time again as being true, but it is not intuitive at all why the world should function this way. The best example I can think of is the Stern-Gerlach experiment (see Sakurai, chapter 1) which I won't discuss until lecture 2, but which makes clear by way of observation this bizarre property of nature even if it's a complete mystery *why* nature works in this way. As I mentioned before it is not the function of quantum mechanics to answer why only how. Weird, I know.

The next fundamental postulate of quantum mechanics relates the probability of the system collapsing to a given eigenket and hence observing the corresponding eigenvalue:

The probability of observing a', which is an eigenvalue of the operator A, of a system in stae $|a\rangle$ is given by $|\langle a'|a\rangle|^2$.

This is consistent with the other postulates we've considered and what we know about physical reality. For example, if $|a\rangle$ is normalized then clearly we have

$$\sum_{a'} |c_{a'}|^2 = 1$$

This makes sense from a probabilistic point of view. If the state must collapse to one of the eigenkets of A then the sum of the probabilities must equal unity. It says that the more character the state $|a\rangle$ has of a given eigenket of A the more likely it is to measure the corresponding eigenvalue of that eigenket. To put this into mathematical terms we need to discuss measurements in more detail.

When making a measurement we typically consider a pure ensemble (We will deal with mixed ensembles when we introduce the density matrix in later lectures)in which we deal with a very large number of identical systems or states. Alternatively, we can think of a measurement on a single system being repeated a very large number of times. These two scenarios are identical from this point-of-view. When we make many measurements of a system in a state $|a\rangle$ we are taking the weighted average of the expansion coefficients with respect to the eigenkets of the observable A. Lets see by way of example why this is so.

Example 3. Suppose we are measuring spin, which we haven't discussed yet in detail, but which we know can only take on the values $+1/2$ and $-1/2$ (here we set $\hbar = 1$, for convenience). An arbitrary ket, say corresponding to a nuclear spin, is in some linear combination of these eigenkets of our observable I_z:

$$|a\rangle = c_1|+1/2\rangle + c_2|-1/2\rangle$$

This means, according to our postulates, that the probability of observing a value +1/2 and -1/2 is given by $|c_1|^2$ and $|c_1|^2$, respectively. However, if we make a single measurement we can only obtain +1/2 or -1/2 and no other value. However what if we make 3 measurements. We might get +1/2,-1/2,+1/2 which on average give 1/6. If we made an infinite number of measurements we would get $1/2 * |c_1|^2 + (-1/2) * |c_2|^2$, exactly, which can take on any value between -1/2 and 1/2.

To make this more mathematical and general we define the expectation value which is akin to an ensemble average as

$$\langle A \rangle \equiv \langle a|A|a\rangle$$

This is easily evaluated to give

$$\langle A \rangle = \sum_{a'} a' |\langle a'|a\rangle|^2$$

Exercise 6. Imagine you had an operator, Q, which is given by

$$Q = a|1\rangle\langle 1| + b|2\rangle\langle 2| + c|1\rangle\langle 2| + d|2\rangle\langle 1|$$

where $|1\rangle$, and $|2\rangle$ are eigenkets of another operator P with eigenvalues p_1 and p_2, respectively. If the system is in a random state with respect to P derive a condition for the coefficients a, b, c, and d that would guarantee that the expectation value of Q would be p_1? ♣

I hope you now have a feel for measurements and their relation to kets, operators, and eigenvalues. The next step is to discuss the relationships between different observables. This topic is extremely important in quantum mechanics. We begin with a definition:

Definition 1.2. The observables A and B are said to be compatible when the corresponding operators commute:

$$[A, B] = 0$$

and incompatible when they don't

$$[A, B] \neq 0$$

where the commutator [,] is defined as

$$[A, B] = AB - BA$$

This definition was needed to state the following theorem:

Theorem 1.3. *If A and B are compatible observables then each eigenket of A is simultaneously an eigenket of B.*

The proof of the compatibility theorem can be found in the supplementary lecture notes or in Sakurai, page 30-31). Let's try to understand the importance of this theorem. If two observables are compatible then as the proof states they share the same eigenket which does not necessarily correspond to the same eigenvalue. So we have

$$A|a'\rangle = a'|a'\rangle$$
$$B|a'\rangle = b'|a'\rangle$$

To be impartial and to represent the fact that both a' and b' are eigenvalues of the same eigenket we write the eigenket as $|a', b'\rangle$. This conveys the fact that for a given state we have can simultaneously obtain the observables A and B.

Example 4. While we haven't talked about angular momentum you may recall that for atomic states we have to supply both the orbital angular momentum squared, $\mathbf{L}^2$, and the z-component of the angular momentum, L_z, to completely define the state. Supplying only one is not sufficient. That is for a given value of l, we can have $m = -l, -l+1, ..., l-1, l$. So it is necessary to supply both l and m to uniquely define the state. The reason, l and m can be simultaneously defined is because $[\mathbf{L}^2, L_z] = 0$, and so $|l, m\rangle$ is an eigenket of both observables. We'll prove this commutator relation later on when we discuss angular momentum and rotations in quantum mechanics.

To make this statement about simultaneous eigenkets more general we need to discuss the possibility of more than two compatible observables. Suppose we could find a maximal set of commuting observables, $A, B, C, ...$ Then the ket $|K'\rangle = |a', b', c', ...\rangle$ would be a simultaneous eigenket of all the observables. In other words the state of the system could be completely described by this ket and furthermore the order in which measurements are made is irrelevent.

Remark. If you're at all familiar with the notion of term symbols in chemistry or group theory than you can understand the applicability of this principle of compatible observables. For example, the atomic term symbol has the form

$$^{2S+1}L_J$$

where $\mathbf{L}$ is the total orbital angular momentum and $\mathbf{S}$ is the total spin angular momentum. $\mathbf{J}$ is simply the sum of two, $\mathbf{J} = \mathbf{L} + \mathbf{S}$. We don't discuss what these terms mean physically, but we can understand the term symbol in terms of compatible observables. Spin and orbital angular momentum are operators in two different spaces, spin space and the infinite-dimensional space spanned by the position eigenkets (physical space). Because these operators operator on different spaces they commute. So the operators, $\mathbf{L}$,$\mathbf{S}$, and $\mathbf{J}$ form a maximal set of commuting observables for atoms, which means that any eigenket of the system can be described by the three quantum numbers, L, S, and J, which are the eigenvalues of the operators $\mathbf{L}$,$\mathbf{S}$, and $\mathbf{J}$. The $2S+1$ comes from degeneracy which we haven't discussed yet, but which is a result of the fact that $[\mathbf{L}^2, \mathbf{L}_z] = 0$. I wish I could spend more time on this fascinating subject, but then the class would go on forever. Sigh...

A strikingly different result occurs when observables are not compatible. In fact, this is what makes quantum mechanics so fascinating and counterintuitive. In fact the phenomenon of nuclear magnetic resonance would not occur were it not for the incompatibility of various operators because we would be unable to effect the nuclear spin. In other words, acting on the spin by an operator that commutes with the previous operator would not change the state of the system and is tantamount to doing nothing. Before getting into the formality of the mathematics needed to generalize things, lets look at a simple example.

Example 5. From linear algebra we know how to represent rotations using matrices. For example, rotation about the z axis by an angle θ can be represented by

$$R_z = \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Similarly rotation about the x axis and y axis are represented by

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos(\theta) & \sin(\theta) \\ 0 & -\sin(\theta) & cos(\theta) \end{pmatrix}$$

$$R_y = \begin{pmatrix} cos(\theta) & 0 & sin(\theta) \\ 0 & 1 & 0 \\ -sin(\theta) & 0 & cos(\theta) \end{pmatrix}$$

If we begin with the vector $\mathbf{r}_i = [1, 0, 0]$ which is along the x-axis and rotate it about the x-axis by 90° and then about the y-axis by 90° we are left with $\mathbf{r}_f = [0, 0, -1]$ which points along -z. However if we first rotated about y and then by x we would get $\mathbf{r}_f = [0, 1, 0]$ which points along y. Clearly the order in which we proceed with rotations is important. The reason is that the first set of operations is given by

$$\mathbf{r}_f = R_{y,90^\circ} R_{x,90^\circ} \mathbf{r}_i$$

while the second set is given by

$$\mathbf{r}_f = R_{x,90^\circ} R_{y,90^\circ} \mathbf{r}_i$$

Since, in general, $[R_x, R_y] \neq 0$ this is not surprising. Later, we will make the connection of rotation matrices with operators and operators with matrices to complete the connection of this example with the physical reality of measurements.

Here's an exercise to further solidify things:

Exercise 7. Prove that, in general, $[R_{x,\theta}, R_{z,\phi}] \neq 0$. Does there exist a ϕ and θ where these two operators do commute? Using Matlab calculate the result of rotation of the vector $[1, 4, -3]$ with $\theta = \pi/3$ and $\phi = \pi/2$ (calculate it for rotation about x and then z and vice-versa). ♣

Now that we have a feel for the consequences of incompatible observables we state without proof one of the most important theorems of quantum mechanics, the uncertainty relation:

Theorem 1.4. *If A and B are observables then for any state the following inequality holds:*

$$\langle(\Delta A)^2\rangle\langle(\Delta B)^2\rangle \geq |\langle[A,B]\rangle|^2/4$$

where $\Delta A \equiv A - \langle A\rangle$ and $\Delta B \equiv B - \langle B\rangle$.

The consequences of this theorem are astounding. ΔA and ΔB can be thought of the uncertainty in A and B, respectively. The uncertainty relation says that if A and B are compatible observables there is no uncertainty in variance of A or B, where the variance is simply $\langle(\Delta A)^2\rangle$ and $\langle(\Delta B)^2\rangle$. That is, both can be known simultaneously with arbitrary precision (limited, of course, by the experimental measurement). However, if the measurements are incompatible, there is necessarily an uncertainty associated with each observable. Furthermore, the less variance in one, the more variance in the other. If one observable is known with perfect certainty, the other is known with perfect uncertainty.

Example 6. The most famous uncertainty relation has to do with position and momentum which we purposefully neglected because they operate on kets in an infinite dimensional (i.e. continuous) Hilbert space. However, it can be shown (see Sakurai, section 1.6), that

$$[x,p] = i\hbar$$

which according to the uncertainty principle implies

$$\langle(\Delta x)^2\rangle\langle(\Delta p)^2\rangle \geq \hbar^2/4$$

or more commonly written as

$$\langle\Delta x\rangle\langle\Delta p\rangle \geq \hbar/2$$

Remark. It is often mistakenly thought that the time/energy uncertainty relation results from the quantum mechanical uncertainty relation written above. Because time has no corresponding operator clearly the uncertainty relation cannot be applied to it. The origin of the time/energy uncertainty relation is completely different and has to do with the Fourier relationship between the two. We'll examine Fourier relationships a bit later in the course.

Let's summarize what we've done up to this point. We made the connection of observables with eigenkets and eigenvalues. That is, all observables are eigenvalues of the corresponding eigenkets of that observable operator. The decision of which eigenket the state collapses onto is determined probabilistically according to the expansion coefficient of the state into the basis kets which are the eigenkets of the given observable operator. The expectation value, or the value we expect to observe in a measurement of a large ensemble of identical systems (i.e. pure ensemble) is given by a weighted average of these expansion coefficient times the eigenvalue. In the case that more than one measurement is taken, say corresponding

to different observables, the order in which the measurements are performed is irrelevant if the observables are compatible. However, if they are incompatible the order is crucial. Furthermore, both observables cannot be measured with absolute certainty simultaneously. A lower bound for the product of variances is given by the uncertainty relation. As a consequence the more precise one observable is known (i.e. smaller variance) the less precise the other observable is known (i.e. larger variance). This is not due to an imperfect method of measurement, rather it is a fundamental property of all systems.

We are now ready to make groundwork into how to actually calculate properties of quantum mechanical systems.

1.4. Matrix Representation

As has already been hinted in previous remarks and examples, operators can be represented in terms of matrices. This greatly facilitates calculations since we can easily manipulate matrices.

Let $\{|a'\rangle\}$ be an orthonormal basis set of kets in Hilbert space. Remember, this means that an ket in the Hilbert space can be written as a linear combination of kets from this basis set. From the completeness relation, any operator, X, can be written as

$$X = \sum_{a''} \sum_{a'} |a'\rangle\langle a'|X|a''\rangle\langle a''|$$

Due to the double sum, there are N^2 terms in this equation which we call matrix elements and arrange as follows:

$$X \doteq \begin{pmatrix} \langle a^{(1)}|X|a^{(1)}\rangle & \langle a^{(1)}|X|a^{(2)}\rangle & . & . & . \\ \langle a^{(2)}|X|a^{(1)}\rangle & \langle a^{(2)}|X|a^{(2)}\rangle & . & . & . \\ . & . & . & & \\ . & . & & . & \\ . & . & & & . \end{pmatrix}$$

The dot over the equal sign is to indicate that this operator is written in this way with respect to a given basis. A matrix by itself has no meaning as an operator. Only when supplied with the basis set does it take on meaning. We now have a recipe for calculating the effect of an operator on an arbitrary ket. Again, we use the completeness relation:

$$X|a\rangle = \sum_{a''} \sum_{a'} |a'\rangle\langle a'|X|a''\rangle\langle a''|a\rangle$$

Recall that $|a\rangle$ can be written in the $\{|a'\rangle\}$ basis

$$|a\rangle = \sum_{a'} |a'\rangle\langle a'|a\rangle$$

which gives

$$|a\rangle \doteq \begin{pmatrix} \langle a^{(1)}|a\rangle \\ \langle a^{(2)}|a\rangle \\ . \\ . \\ . \end{pmatrix}$$

As before this representation is meaningful only when we specify the basis set. Therefore, to calculate the effect of X on $|a\rangle$ we simply multiply the matrix by the vector in the usual way:

$$X|a\rangle \doteq \begin{pmatrix} \langle a^{(1)}|X|a^{(1)}\rangle & \langle a^{(1)}|X|a^{(2)}\rangle & . & . & . \\ \langle a^{(2)}|X|a^{(1)}\rangle & \langle a^{(2)}|X|a^{(2)}\rangle & . & . & . \\ . & . & . & & \\ . & . & & . & \\ . & . & & & . \end{pmatrix} \begin{pmatrix} \langle a^{(1)}|a\rangle \\ \langle a^{(2)}|a\rangle \\ . \\ . \\ . \end{pmatrix}$$

A special case of matrix representation that's important to consider is when the eigenkets themselves are used. In this case, the observable A is represented by a matrix with the eigenvalues as diagonal elements and 0 for the off-diagonal elements.

$$\langle a'|A|a'\rangle = a'\langle a'|a'\rangle = a'$$

and

$$\langle a'|A|a''\rangle = a'\langle a'|a''\rangle = 0$$

Exercise 8. Write down the matrix representation for the operator Q from exercise 6 in the P basis and in the Q basis. What are the observables of Q? ♣

Often, it's desirable to change from one basis to another when dealing with incompatible observables. Later on we'll see examples when this is useful - where careful choice of a basis greatly simplifies evaluation of the dynamics of the system. In general, the idea of applying transformations is critical for simplifying a problem. For example, when dealing with spherically symmetric systems, it's very inconvenient to work in Cartesian coordinates. Similarly, in quantum mechanics changing basis can greatly simplify the analysis. Next we'll present how this is done formally and leave it for later to apply this method to specific systems of interest.

Let A and B be incompatible observables with eigkets $\{|a'\rangle\}$ and $\{|b'\rangle\}$, respectively. If we worked in the $\{|a'\rangle\}$ basis, A would be represented by a diagonal matrix. We're interested in finding a transformation that allows us to move from a representation in $\{|a'\rangle\}$ to one in $\{|b'\rangle\}$. Mathematically, we want an operator, U, that satisfies

$$|b^{(i)}\rangle = U|a^{(i)}\rangle$$

Since $\{|a'\rangle\}$ consists of orthonormal kets U is simple to construct

$$U = \sum_k |b^{(k)}\rangle\langle a^{(k)}|$$

Exercise 9. Show that U is unitary. That is, $U^\dagger U = UU^\dagger = 1$. Furthermore, show that all unitary matrices satisfy $|det(U)| = 1$. ♣

Now, we need to write the matrix elements of A in the $\{|b'\rangle\}$ basis in terms of U and matrix elements of A in the $\{|a'\rangle\}$ basis.

$$\langle b^{(k)}|A|b^{(l)}\rangle = \sum_m \sum_n \langle b^{(k)}|a^{(m)}\rangle\langle a^{(m)}|A|a^{(n)}\rangle\langle a^{(n)}|b^{(l)}\rangle$$

Substituting, $|b^{(i)}\rangle = U|a^{(i)}\rangle$ with $i = k$ (and complex conjugating)and $i = l$ gives

$$\langle b^{(k)}|A|b^{(l)}\rangle = \sum_m \sum_n \langle a^{(k)}|U^\dagger|a^{(m)}\rangle\langle a^{(m)}|A|a^{(n)}\rangle\langle a^{(n)}|U|a^{(l)}\rangle$$

In matrix notation this is simply

$$A_{b'} = U^\dagger A_{a'} U$$

where the subscript indicates the basis used for the matrix representation of the operator A.

Remark. Of course this is exactly the result you obtain in ordinary linear algebra when moving from one basis to another. The formalism of Dirac combined with the notion of matrix representation makes all the tools of linear algebra available to quantum mechanics. From now on we'll just list important properties of operators and their matrix representations instead of proving them using bras, kets, and the completeness relation. Because of the correspondence with linear algebra and results we know about matrices we can be confident about applying those same rules to operators.

An extremely important notion is that of trace which is defined as the sum of the diagonal elements of a matrix. The trace has some important properties which will be commented on soon. The two that most concern us are the following

$$tr(AB) = tr(BA)$$
$$tr(U^\dagger AU) = tr(A)$$

The second identity warrants some discussion (the first is simply very useful for calculations). It says that the trace is invariant to choice of basis. That is the trace of an operator is an important measure of that operator because it's independent of representation. This hints that the trace is somehow connected to the observable and the measurement, which should, of course, be independent of the basis chosen to represent it. We'll see this in lecture 2.

Exercise 10. Show that $A_{a'}$ and $A_{b'}$ share the same set of eigenvalues. ♣

To motivate further the concept of matrix transformations such as a change-of-basis transformation described above, consider a case in which you work in one basis for a given system. Of course, you can choose a basis so that the observable is diagonal by simply taking the set of eigenkets of the operator corresponding to that observable. As previously mentioned, the diagonal elements will be the eigenvalues of these eigenkets. Now suppose that the system changed a bit so that now the eigenkets of the previous system were no longer eigenkets of the new, perturbed system. Then the matrix representation would no longer be diagonal. To diagonalize the new matrix we need to find its eigenkets and eigenvalues. The unitary matrix that does this is the one discussed above. More concrete examples will make this even more clear as we put the mathematics into practice. Again, we hold off on doing this until we discuss Hamiltonians in lecture 2.

Exercise 11. Here are some problems for those a bit rusty on their linear algebra skills:

1. Diagonalize the following matrix using Matlab:

$$A \doteq \begin{pmatrix} 1 & 2-i & 0 \\ 2+i & 3 & i \\ 0 & -i & 5 \end{pmatrix}$$

By diagonalize I mean find a unitary matrix U such that $U^\dagger AU$ is diagonal. Is A Hermitian? If it were not, would U be necessarily unitary?

2. What is the matrix representation of A in the basis $[\ 2 \quad 1 \quad 0\ ]$, $[\ 1 \quad 4 \quad 3\ ]$, and $[\ 1 \quad 0 \quad -1\ ]$, assuming that A in 1. is written in the $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ basis set (i.e. the standard basis).

1.5. Summary of Postulates

- The state of a quantum system is represented by a state vector in a Hilbert space which contains all the information about the physical state of the system.
- Dynamical variables are represented by Hermitian operators.
- A measuremnt always causes the system to jump into an eigenstate of the dynamical variable that is being measured. This means that a measurement of by an operator, A, results in the eigenvalues of A.
- The probability of observing a', which is an eigenvalue of the operator A, of a system in state $|\alpha\rangle$ is given by $|\langle a'|\alpha\rangle|^2$, the square of the expansion coefficient.

1.6. Bibliography

- J.J.Sakurai "Modern Quantum Mechanics (Revised Edition)", Addison-Wesley Publishing (1994)
- R.E. Larson, B.H. Edwards "Elementary Linear Algebra (Third Edition)", D.C. Heath and Company (1996)

- R. Zwanzig "Nonequilibrium Statistical Mechanics", Oxford University Press (2001)
- M. Goldman "Quantum Description of High-Resolution NMR in Liquids", Oxford University Press (1998)