FOUNDATIONS OF QUANTUM THEORY: NON-RELATIVISTIC APPROACH

Dr. Sandeep K. Goyal
Department of Physical Sciences
IISER Mohali

Week-01
Lecture-01

Axiomatic Approach to Quantum Mechanics - Part 01

Hello everyone. In this lecture, we will talk about the axiomatic approach to quantum mechanics. Quantum mechanics is one of those few theories in physics which have their foundations on a number of axioms. More specifically, there are four axioms that describes the theory of quantum mechanics. So we have quantum in the center.

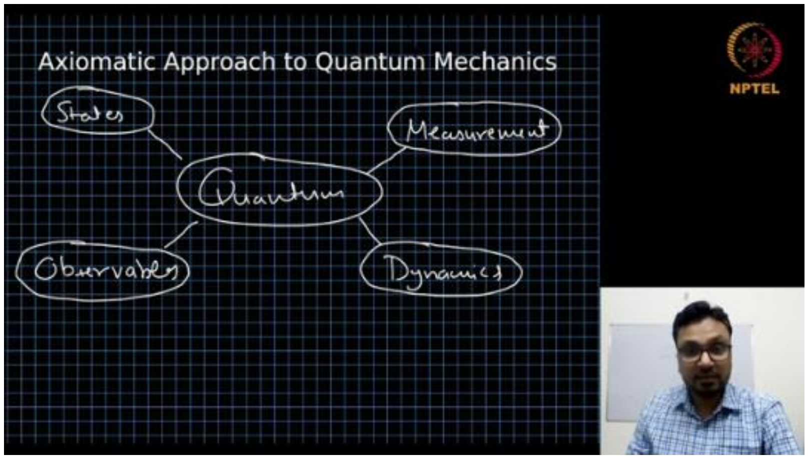
And this depends on four axioms. First axiom describes the state. So in this axiom, we discuss what can be called a state of a quantum system, what are its properties, what can be done with the state, what cannot be done with the state and stuff like that. The second axiom depends on the observable, that is, what is the mathematical structure of observables? What is their relation with the states?

What is their relation with the measurement operators? What is their relationship with the dynamics of the system? And why certain mathematical structures can be called observables? And what kind of mathematical structures can be called observables? All those things have been attributed to the observable axiom.

Third is the dynamics. This axiom talks about how a quantum system evolved, what is the dynamics of the quantum system and what equation can describe it. So, in this we will talk about Schrodinger equation and how we arrive at the Schrodinger equation. And fourth axiom, last but not the least is the measurement axiom. So in this we will discuss what do we mean by a measurement on a quantum system, what are the outcomes of that measurement, what happened to the system after measurement and how this measurement is related to the observables.

So all these things all these four axioms together contribute to quantum mechanics. If we understand these four axioms at its foundational level, then the quantum mechanics will not be as mysterious as everyone think it is. These four axioms is everything which is there to be offered by quantum mechanics. If we understand these axioms, there is nothing else to be understood. Everything can be derived from the understanding of these four axioms. The first one is about the states that how we represent the state of the quantum system and what are the things we have to keep in mind, what is the mathematical background of the states, state space and stuff like that.

(Refer slide time: 2:48)



The observable axiom tells us what mathematical structure can be called observables and why it is so, why that particular mathematical structure is called observable and what are the properties of observables ? The third one is about the dynamics, how we can evolve a quantum state in time, how the quantum system evolves in time. How do we perform transformation from one quantum state to other quantum state and what are the physical processes and all those are related topics but they all are subset of the dynamical dynamics postulate for axiom. And fourth one, which is the measurement postulate, and this tells us that what kind of measurements are allowed on the system, what happened to the system after measurement, what are the outcomes of the measurement and how we can find the measurable quantity in the lab using this measurement postulate and the observable postulate.

So, we will start with the state. Quantum system is a physical system. So a mathematical structure which can contain every possible information about the quantum system should be called a state, can be a valid state. Let me repeat the statement, we have a quantum system and let us say we have a hypothetical mathematical structure and let us call it Ψ and we do not know what this is at the moment. It can be a function, it can be a operator, it can be a vector, it can be a complex number, it can be anything. Right now, we do not know what it is.

We only know that if we use it in a specific manner, in some convenient manner, then we can retrieve all the information there is to know about a quantum state, quantum system.

Okay so what I mean by this is, if we are allowed certain measurements on the system like position measurement, momentum measurement, temperature and whatever measurement we can think of in the lab on a quantum system, if we perform those measurements, we will get certain outcomes in the measurement. Those quantities we should be able to retrieve from the mathematical structure and if we can find such Ψ then we will call it the state of the quantum system. So, Ψ is the state of the quantum state system if it contains all the information, all the information about a given quantum system. So, of course, we can always say that there is too much information in any given quantum system.

So how can any mathematical structure contain all that information? We will talk about that soon when we talk about other postulates and little bit more details of the states. But this is what we will base our whole course on that our state is a mathematical structure which contains all the information about the quantum system. But there is a small catch here. When we say all the information, we only mean that the information that can be retrieved from an experiment or from all the possible experiments. So, if quantum mechanics does not allow for certain information to be known, that information will not be contained in the state set.

But those are slightly more advanced topic we will discuss them later on when we talk about other topics like the completely positive maps and state tomography and state preparation methods. When we talk about those things, we will know what we mean by what I just said that there is certain information which is not knowable from the experiments and that information will not be contained in the state side. So, mathematically, whenever we write the state of a quantum system, we put this symbol. So, whenever we see this symbol, it will mean that it is a quantum state and it is called ket. So, $|\Psi\rangle$ means the quantum state Ψ. And if we take the adjoint of this $(|\Psi\rangle)^\dagger$, it will be this symbol and that is bra psi, $\langle\Psi|$.

In the traditional quantum mechanics, the state $|\Psi\rangle$ is an element from the Hilbert space H. And this Hilbert space is the space which contains all the states of the quantum system.

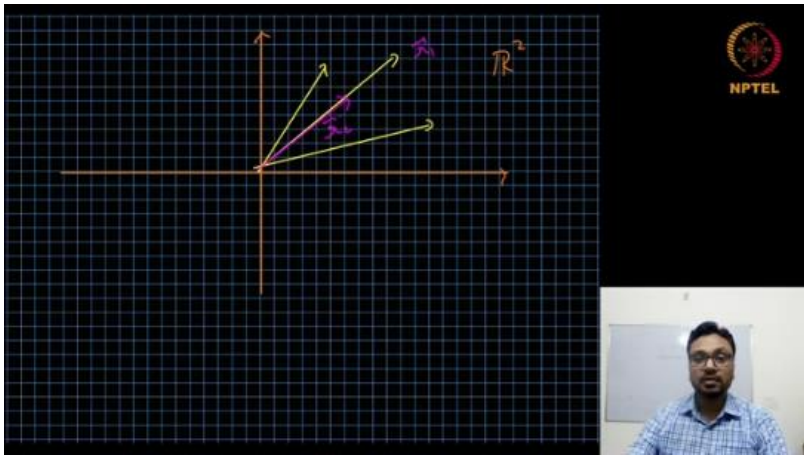
Let us say S, quantum system is S. So, we put subscript S here just to represent that this Hilbert space belongs to the quantum system S. So, a state of the quantum system $|\Psi\rangle$ belongs to the Hilbert space Hs and all the states of the quantum system belong to this Hilbert space. So, that is our first mathematical definition of the state that it is a vector in the Hilbert space Hs. Let me remind you that Hilbert space is a linear vector space over which we have defined the inner product and the metric of length distance. $|\Psi\rangle$ is also called pure state as opposed to mixed state which we will be discussing later on, not in this lecture. The properties of this state $|\Psi\rangle$ is that $|\Psi\rangle$ is normalized, that the inner product of $|\Psi\rangle$ with itself is 1. Notice that if we have two states $|\Psi_1\rangle$ and $|\Psi_2\rangle$ they are the valid state and they are they belong to the Hilbert space Hs, then their superposition, $\alpha\,|\Psi_1\rangle + \beta|\Psi_2\rangle$, which also belongs to the Hilbert space because that's the property of the Hilbert space and that is also a valid state as long as $|\alpha|^2 + |\beta|^2 = 1$.

This normalization condition is very important that the state should be normalized and we can take the superposition of any arbitrary number of states, the $\sum_n \alpha_n\,|\Psi_\mathrm{n}\rangle$ is also a valid state as long as $\sum_n |\alpha_n|^2 = 1$, where α_n are the complex numbers. So, these are the properties of the states here that they should be normalized. There is a proposition as long as it is normalized is a valid state and they are vectors in the Hilbert space Hs. Although the state $|\Psi\rangle$ is a vector in the Hilbert space Hs, not all the vectors in the Hilbert space are states. In fact, two states, two vectors from the Hilbert space $|\phi\rangle$ and $|\Psi\rangle$ are equivalent if or they represent the same quantum state if $|\Psi\rangle$ is some scalar lambda times $|\phi\rangle$. So, it means the set of all the states $|\phi\rangle$ such that $\lambda|\phi\rangle = |\Psi\rangle$ represent the same state. It means we can divide the vectors of Hilbert space into sets of vectors which represent the same state. Like the way we defined here that if a vector, if two vectors are different only by a scalar factor, then they represent the same state. So, in that way Hilbert space can be divided into set of such vectors.

To understand what we are doing here, let us consider a two-dimensional Hilbert space. Two-dimensional real space, we have y axis, we have x axis. So, this is R^2, this is a two-dimensional vector space and we can define the inner product as the Cartesian product. Then we have many vectors here. For example, we can take few vectors 1 and 2 and 3. These are different vectors. But now consider another vector. One is this vector, let us call it $\mathbf{r}_1$ vector. And we take another vector along this line, but up to a smaller distance, $\mathbf{r}_2$ vector.

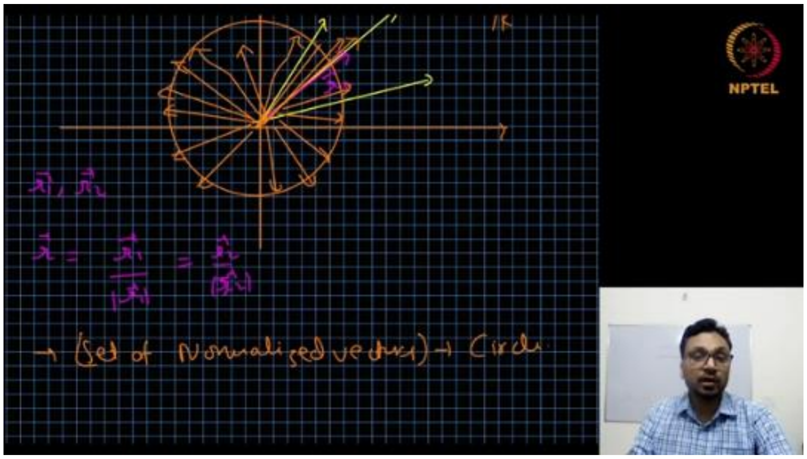
So there are two vectors $\mathbf{r}_1$ and $\mathbf{r}_2$. They are only different in the size. So what we can do is we can find a vector $\mathbf{r}$ which is $\mathbf{r_1}/|\mathbf{r_1}|$. So it's a unit vector and it is $\mathbf{r_2}/|\mathbf{r_2}|$. And you can find many such vectors which will if we normalize it will become $\mathbf{r}$ vector. So, we take that normalized vector. We take all the normalized vectors in this space and each normalized vector will represent one quantum state. and in R^2 a set of normalized vectors, if we take a set of normalized vector that will make a circle, because if we take the tips and if we join them, we get a circle.

(Refer slide time: 12:21)



So this circle in the whole R^2 space represents the quantum state but this was a hypothetical example of R because it's easy to visualize. Now think of a Hilbert space which is a complex vector space which will be very hard to imagine but just think of it as a two-dimensional R^2 as a as a specific version of the Hilbert space in that we take all the vectors which are only different in the norm and we identify it as one vector. So, one normalized vector will be the representative vector of that. We find all such different vectors and they all together will form a circle in R^2 and they will form a more complicated mathematical structure in a more complicated Hilbert space, and that is called the complex projective Hilbert space. So, we identify all the vectors which are different only in the norm as one vector and we normalize it and we take the collection of all the vectors all the normalized vectors.

(Refer slide time: 13:48)



This is the complex projective Hilbert space for Hs, like we showed that the circle is the complex projective space for the R^2. This is also called a ray space. And ray space is the space which contains all the quantum states of a given system, and all the points in a ray space are valid quantum states unlike the Hilbert space where all the vectors do not represent quantum states, but in Ray space all the vectors represent quantum states.

Next is observable. What are the observables for quantum mechanics is the question we will be addressing in this axiom. The experience, although the states being a part of Hilbert space can contain complex numbers, but whenever we observe something, whenever we measure something, the outcome in a lab will be always real.

So observable is some mathematical structure, let us call it Z, some mathematical structure such that the outcomes are real, and, the outcome should be distinguishable. They should be independent. In mathematical terms, it will be orthogonal. So, we are looking for a mathematical object, mathematical structure such that it has some quantity which are real numbers and some quantity which are some mathematical structure which are orthogonal. So, one candidate for that are the Hermitian operators, or self-adjoint operator. So, if we have a self-adjoint operator Z, then we know it is $Z = Z^\dagger$, that is the definition of self-adjoint or Hermitian operator.

We know that if the operator is self-adjoint, then the eigenvalues are real and eigenvector are orthogonal. So, we can identify the eigenvalue as the outcomes in lab. So, whenever we measure this observable, then the eigenvalues will be the outcomes of the measurement and eigenvector will be the state we will be getting. So, in that way the Hermitian operators satisfy the requirement posed by the physical consideration. So, in quantum mechanics we will consider Hermitian operators as the observable of the system. If we have a Hilbert space Hs and the dimension of this Hilbert space is d. So, the

Hilbert system we are considering is a d dimensional system. So, this can be a physical example of this can be an atom where we are considering only d energy levels of that atom.

So, one such quantum system will have dimension d or spin half electron then it has only two quantum states. So, d becomes two. So, if we have a Hilbert space Hs, where the dimension of the Hilbert space is d, then the observables or the Hermitian operators Z will be d by d matrices. And they belong to a set B acting on Hs. So, this B is a set of operators acting on the Hilbert space H. So, the operators which are allowed to act on the state or on the vectors of the Hilbert space H. We are representing that with B of Hs. So, the Hermitian operator Z belongs to this set. Not just set, it belongs to, this is general set, so it belongs to the set of Hermitian operators.

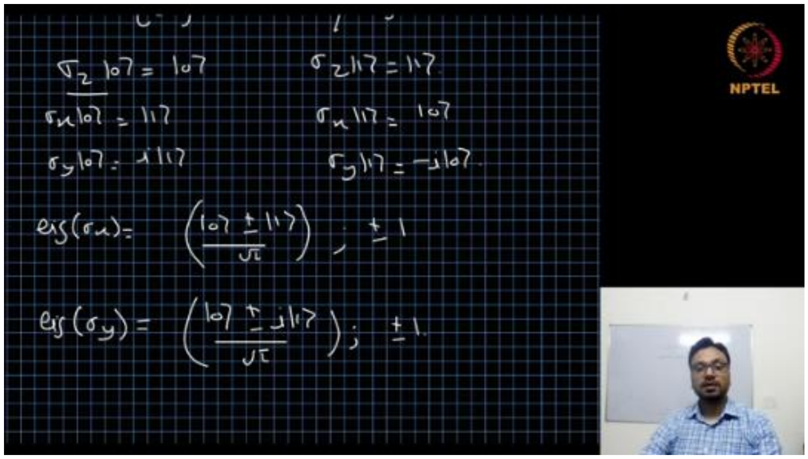
The Hermitian operators form a vector space. The set of Hermitian operator, set of all the Hermitian operator form a vector space over real field. The dimension of this vector space is d^2. So, it means there are d^2 mutually orthogonal and independent Hermitian operators you can find in this vector space. For example, for d equals two, we can have a basis for this vector space and this is the poly basis. I am giving one example. There are of course more than one example for it. So, the poly basis have four elements. One is identity, is a 2 by 2 identity. Then we have sigma x (σ_x), sigma y (σ_y) and sigma z (σ_z), where identity is $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, σ_x is $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. σ_y is $\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$. The element at σ_z is $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. So here,

identity is the trivial operator and the eigenvalues of that are 1 and 1.

σ_x has eigenvalue 1 and -1, σ_y has eigenvalue 1 and -1, σ_z has eigenvalue 1 and -1. If we represent $|0\rangle$ *as* $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ vector and $|1\rangle$ as $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ vector, then we can

see σ_z acting on $|0\rangle$ is just $|0\rangle$ and σ_z acting on $|1\rangle$ is just $|1\rangle$. So, in that way, we can see σ_x acting on $|0\rangle$ will give us $|1\rangle$, σ_x acting on $|1\rangle$ will give us $|0\rangle$, σ_y acting on $|0\rangle$ will give us i times $|1\rangle$ and σ_y acting on $|1\rangle$ will give us -i times $|0\rangle$. Not just that, so $|0\rangle$ and $|1\rangle$ are the eigenvectors of σ_z, eigenvector of σ_x can be written as $(|0\rangle \pm |1\rangle)/\sqrt{2}$ and these are of course two eigenvalues ± 1, similarly eigenvectors of σ_y are $(|0\rangle \pm i|1\rangle)/\sqrt{2}$ and they correspond to the ± 1 eigenvalues. Identity operator is the trivial operator, every vector is an eigenvector of identity. So, we do not need to mention that here.

(Refer slide time: 22:52)



Since the poly basis is a basis for 2 by 2 Hermitian operator, any Hermitian matrix H can be written as $\sum_{\mu=0}^{3} h_\mu \, \sigma_\mu$, where h_μ are real numbers, and σ_0 is identity, σ_1 is sigma x, σ_2 is sigma y and σ_3 is sigma z. Although it is not required, but we can also define the inner product over the vector space of the Hermitian operator. So, the inner product between two matrices A and B, we can define as trace of $A^\dagger B$, so for Hermitian operator it will be just trace of AB. This is how we can define the inner product over the operators. Next is the dynamics or postulate dynamics of quantum system. So, in this postulate we discuss how a quantum system evolves in time. So, if we are given a quantum system in certain initial state, what is the state of the system after time t ?

And this is where the Schrodinger equation comes into the play. Schrodinger equation is what defines the dynamics of the quantum system. So, how do we arrive at the Schrodinger equation? So, first we notice that quantum mechanics is the wave theory. It is inspired from the wave equation.

And the general solution of a wave equation is $e^{i(\boldsymbol{k}.\boldsymbol{r}-\omega t)}$, where **k** is the wave vector, which tells us the direction of propagation and the wavelength of the light, **r** is the position, ω is the frequency, and t is the time. This is, so any solution of the wave equation will be proportional to this. Now, we see that in quantum mechanics, the first thing we know about quantum mechanics is the energy is quantized. So, energy is always $\hbar\omega$ where $\hbar$ is $h/2\pi$, where h is the Planck's constant. So, for a given frequency, if we have a wave of given frequency ω, then the energy of that wave is quantized, it cannot be any arbitrary value, it has to be an integer multiple of this $\hbar\omega$.

So, it means if we have a state $|\Psi\rangle$ of the quantum system, and then we have some operator, some energy operator Ê, that should give us $\hbar\omega$ times the state $|\Psi\rangle$. This is the meaning of this condition that if energy is $\hbar\omega$, so $\hbar\omega$ is the eigenvalue of the state of the operator, energy operator. And if we have the state like the wave state, wave equation, solution of the wave equation, then the operator acting on this state should give us $\hbar\omega$ times the state side. So, if we compare, if we say the $|\Psi\rangle$ is proportional to $e^{i(\boldsymbol{k}.\boldsymbol{r}-\omega t)}$, from here we can see that if we have $i\hbar\frac{d}{dt}$ of $|\Psi\rangle$, we get $\hbar\omega$ times $|\Psi\rangle$, the same state $|\Psi\rangle$. So, in that way, it seems that the energy can be represented, energy operator can be represented by $i\hbar\frac{d}{dt}$. The time derivative is proportional to the energy operator.

The second quantization or not second quantization, the second thing in quantum mechanics is that the momentum is quantized. So, the **p** is $\hbar\mathbf{k}$. If the state of the quantum particle is represented by this expression, where **k** is the wave vector, then the momentum or the momentum operator or momentum vector of the particle is proportional to the **k** vector or the momentum is inversely proportional to the wavelength. So, if this equation has to be satisfied, then we can see that $-i\hbar\nabla$ operator acting on $|\Psi\rangle$ will give us $\hbar\mathbf{k}|\Psi\rangle$. We can substitute $|\Psi\rangle$ from this equation and we can see that this is indeed true.

So, what we have established is that $i\hbar\frac{d}{dt}$ is equivalent to the energy operator

and -$i\hbar\nabla$ operator is proportional to the momentum operator. Now, we know energy is $\frac{p^2}{2m}+V(r)$ where $\frac{p^2}{2m}$ is the kinetic energy and V(r) is the potential energy. Now, if we substitute it here, then we get $i\hbar\frac{d}{dt}\,|\Psi\rangle$ will be equal to $\frac{-\hbar^2}{2m}\nabla^2|\Psi\rangle+V(r)\,|\Psi\rangle$. So, by arguing the energy quantization, momentum

quantization, the solution of the wave equation and the relation of the quantum mechanics with the wave equation, we have arrived at an equation in which on the left-hand side we have time derivative and on the right hand side we have potential energy and kinetic energy. And this is called the Schrodinger equation, and this equation characterizes the evolution of dynamics of a quantum system.

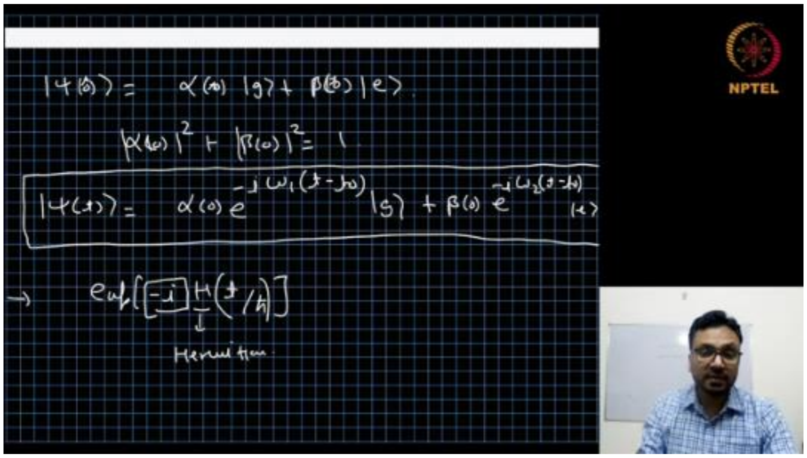
Further, we can see that $\frac{p^2}{2m}+V(r),$ that is kinetic energy plus potential energy is H, the

Hamiltonian of the system. So, we have a quantum system and Hamiltonian is defined over that system. Hamiltonian for an isolated system is just total energy. So, the operator H is the Hamiltonian operator and the Schrodinger equation can be written as $i\hbar\frac{d}{dt}\,|\Psi\rangle = H\,|\Psi\rangle$. And for an isolated system, H is time independent. So, we can write formally the solution $|\Psi(t)\rangle$ at time t equals $\exp[-i\, H\, (t - t_0)/\hbar]\, |\Psi(t_0)\rangle$.

Since H is Hamiltonian, it's an observable of the system so it must be Hermitian. It means H has a spectral decomposition where we have $H = \sum_{n=1}^{d} \lambda_n |\Psi_n\rangle\langle\Psi_n|$. λ_n and $|\Psi_n\rangle$ are the eigenvalue and eigenvectors of the Hamiltonian where n is from 1 to d, where d is the dimension. If for the case when d is 2, then we have Hamiltonian, which is $H = \lambda_1|\Psi_1\rangle\langle\Psi_1| + \lambda_2|\Psi_2\rangle\langle\Psi_2|$. Or we can just say lambda 1 or we can say E_1 for the energy. Let me put it here also E_1, ground state, plus E_2, the excited state ($E_1|g\rangle\langle g| + E_2|e\rangle\langle e|$). There are only two levels of the quantum system. So, one is ground state and other is excited state.

So, we are assuming E_1 is smaller than E_2, so that we can call it ground state. Now, in that case, if that is the case, then $\exp[-iH(t-t_0)/\hbar]$ can be written as since $|g\rangle$ and $|e\rangle$ are orthogonal to each other this thing can be written as exponential of $-i\, E_1(t-t_0)/\hbar$ or I would like to write E_1 as $\hbar\omega_1$ and E_2 as $\hbar\omega_2$ so that we can cancel $\hbar$ when we get $\exp[-i\,\omega_1(t-t_0)]\, |g\rangle\langle g| + \exp[-i\,\omega_2(t-t_0)]\, |e\rangle\langle e|$. What is the benefit of doing this? Now, we have a state $|\Psi(0)\rangle$ which also since it is a two-dimensional state and $|g\rangle$ and $|e\rangle$ are orthonormal basis we can write $|\Psi(0)\rangle = \alpha(0)|g\rangle + \beta(0)|e\rangle$, where $\alpha(0)$ and $\beta(0)$ are the complex coefficients, such that $|\alpha(0)|^2 + |\beta(0)|^2 = 1$. Once we have this, we can easily write $|\Psi(t)\rangle$ as $\alpha(0)\exp[-i\,\omega_1(t-t_0)]\,|g\rangle + \beta(0)\exp[-i\,\omega_2(t-t_0)]\,|e\rangle$.

(Refer slide time: 35:45)



And this is how we can find the evolution of the quantum system. So what we did is we have a Schrodinger equation in which the time derivative of a state is proportional to the Hamiltonian acting on the same state. From here we can see the Hamiltonian is the generator of the time evolution in the quantum system. So, if we are given a Hamiltonian, we can find the evolution of any arbitrary state by decomposing the state into the eigen basis of the Hamiltonian. And in that way, we can proceed with the evolution. Now, one thing to notice here is the evolution was done by an operator exp[-iHt/$\hbar$]. Here we have scalar t/$\hbar$, we have another scalar -i and we have a Hermitian operator. So, i times Hermitian or -i times Hermitian is an anti-Hermitian operator, exponential of an anti-Hermitian operator is a unitary operator and this unitary operator is taking one state to other state.

It so happened that this unitary operator is the time evolution operator. So, the state evolved in time. But in general, t can be any arbitrary parameter, H can be any arbitrary Hermitian operator and there is the i. So, it means unitary transformation, unitary operators transform a state to another state, $|\Psi_1\rangle$ to $|\Psi_2\rangle$. So, in general, any transformation from $|\Psi_1\rangle$ to $|\Psi_2\rangle$ is governed by a unitary operator. If we see some transformation happening on a quantum system, which is not unitary, then we can be sure that we do not have the full information about the system. There is another hidden part of the system which we cannot see. But if we include that part, then the whole dynamics will look unitary.

The whole transformation will look unitary. In that way, the dynamics and the transformations are governed by unitary operators.