Let us resume our formal study of quantum mechanics by first asking if there are any questions and that will give us a starting point. Last time we disused the uncertainty principle a little bit. We tried to point out that quantum mechanics has, by its very nature a certain degree of uncertainty in our ability to completely understand or completely specify numerical values for physical observables.

Now of course we must quantify this and let me say that only part of the statement is right where we often tend to say that quantum mechanics is all about probabilities and that it's indeterminate and it's uncertain. The rules for calculating probabilities are completely known. They are deterministic. There is nothing uncertain about that. But the fact is quantum mechanics does say that associated with physical measurable quantities; we have a certain intrinsic statistical nature which means that, you can only talk in terms of probability distributions and all the other things associated with probability distributions like average values, mean square values, standard deviations and so on. So in that sense, it differs from classical Hamiltonian mechanics for instance, where we could specify the state of a system by a very precise point and phase space. The idea of a point in phase space is lost once you go to quantum mechanics and that's the way things are. Now the frame work in which you discuss a quantum mechanism is mathematical in its most elementary form.

It's more elementary than what you need for Hamiltonian mechanics. We need to have fairly sophisticated concepts such as Poisson brackets, symplectic structure and so on. In quantum mechanics, this is replaced by a very linear kind of a structure where the states of systems are specified not by points in some phase space but by an element of a linear vector space. This element is called the state of a system and it's an abstract concept. The idea is that there exists something called the state of a system which happens to be a member of a linear vector space. This state of the system potentially can give you all the information you can find about the system. Just as in classical mechanics, once I tell you that the system has n degrees of freedom, you associate with it a 2n dimensional phase space, you find the generalized coordinates of momenta. A point in that phase space tells you the state of the system specifying all the q's and p's .you have to throw that out in quantum mechanics.

Now the experiments which led us to this kind of description, over the years grew and in the early days, the interpretation of quantum mechanics posed a very severe problem. The formulation of quantum mechanics itself posed a fairly deep problem but eventually things got ironed out and by 1926 or so, the pre-formalism was replaced.

Since that time, many layers have been added to it. But the original foundational formalism, that was initiated in a very remarkable short time in the early 1920's still remains in place.

(Refer Slide Time: 00:06:22 min)
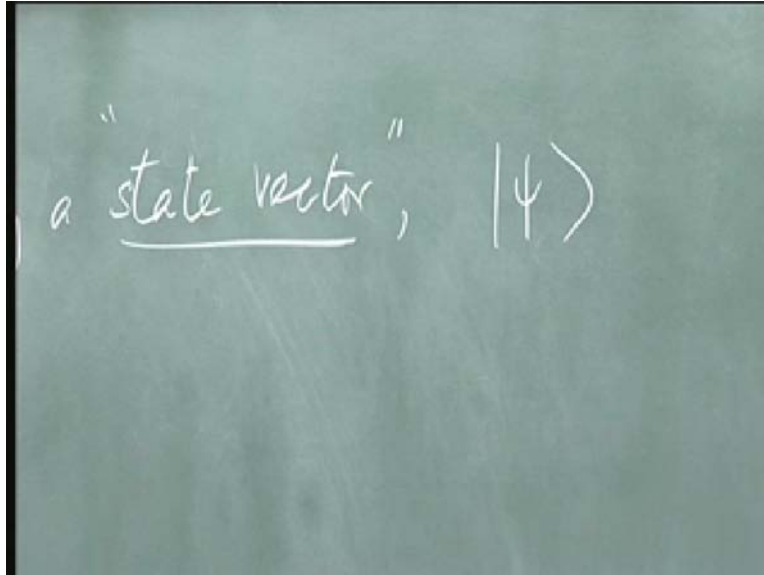
So let me start by saying that a quantum mechanical system is described or specified by a state vector. I use the word vector because it's an element of a linear vector space and a kind of symbols we would use for it changes from different books and authors. But now there are standard symbols for these states. The one that I am going to use is called Dirac notation and I will say lot about Dirac notation as we go along. It is denoted by something like that (Refer Slide Time: 00:06:22).

I will use Greek letters for these state vectors and I am going to put these mysterious angular brackets here just to tell you that it is an element of a liner vector space. Now whenever I say linear vector space, you imagine a three dimensional Euclidean space, for example. The three dimensional Euclidean space is a linear vector space. Therefore all the properties that we are going to talk about for state vectors can be imagined by thinking about ordinary three dimensional vectors and using vector algebra to add them. Now this state vector is a function of time. Let me write psi of t with a capital psi. I will use capital for the moment.

(Refer Slide Time: 00:07:13)
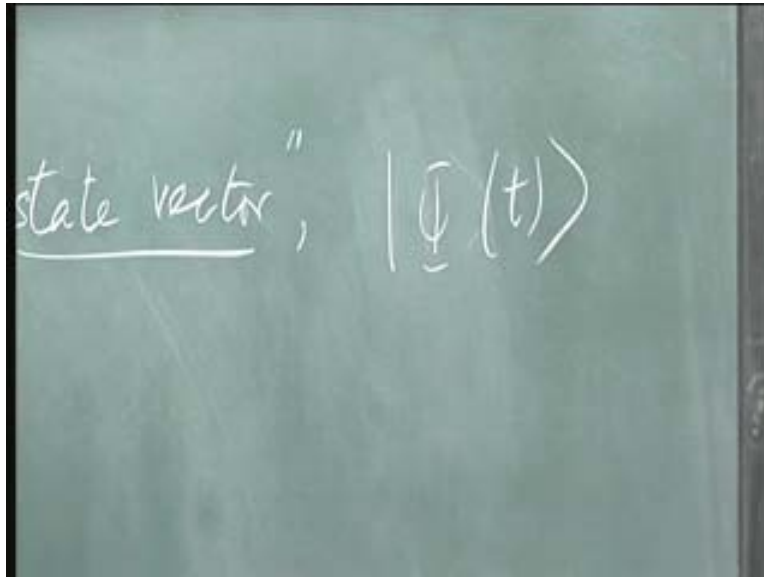
We will see why I use a capital letter because I want to use a small psi for something else. In order to find out what happens to the state as time goes along, we have to prescribe a rule of evolution just as in classical physics we prescribed Hamilton's equations of motion. In exactly the same way, I am going to prescribe a rule for this psi of t and afterwards we will come back and simultaneously we will interpret what the psi of t and how it gives you information on various quantities and so on. But first a little bit of a digression on linear vector spaces.
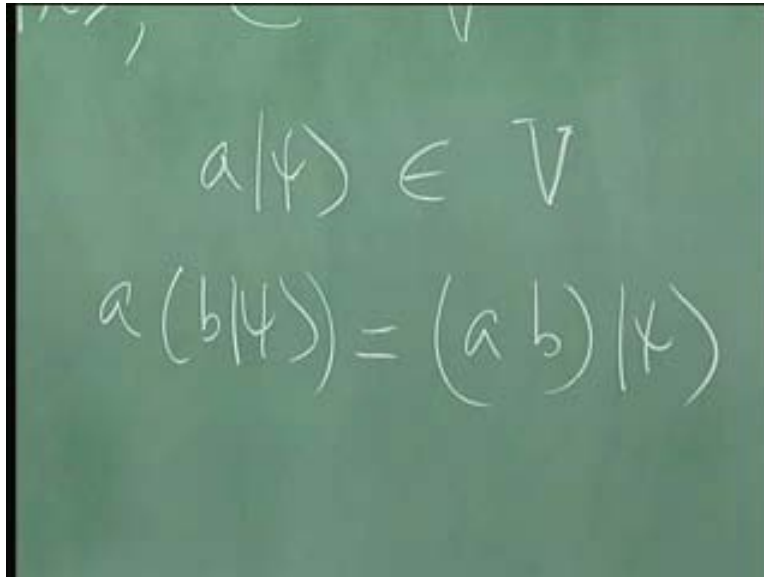
(Refer Slide Time: 00:08:07 min)
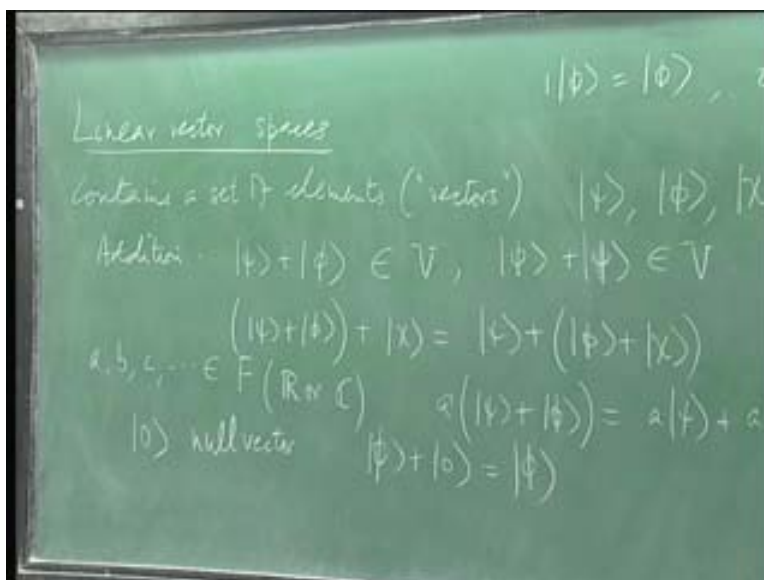
This is essential because I will assume that you know about linear vector spaces. So let me go through a small mathematical digression on linear vector spaces. Essentially if you know about matrices you already know all the mathematics you need. So let me reassure you that there is not much left and all you need to know is how to handle matrices and that too square matrix. Now let me formally define a linear vector space. I will leave out a few things here and there but we will fill them in as we go along. So this contains a set of elements which we called vectors and I need a notation for these elements. Let's call this notation as psi, phi, chi etc. I use Greek letters for these vectors. These are called ket vectors for a reason which I'll explain subsequently. What I use is a funny angle or bracket for it; just like tensor notation in ordinary you can calculate where we have the summation convention and the index notation.

So it helps you do calculation very easily once you use this notation. In fact half the battle is won if you express things in the right notation. And this is exactly what the Dirac notation does. So it contains a set of elements among which you define certain operations. The fundamental one is that of addition. So you say that you add two vectors. Linear vector space V contains a set of elements. So it says that if psi and phi are elements of V, then psi + phi is also an element of V. So you add two vectors and you get another vector which also belongs to the same space. Moreover this addition is associative, in the sense that you could add the vectors in any order you like. It is also true that this addition is commutative. You can add them in either order.

Importantly, the liner vector space is set of elements or vectors defined in a certain field and this field is a field of scalars. So you simultaneously introduce a set of numbers or scalars which belong to a field. We don't want to get into what a field is right now. The real numbers and complex numbers form a field. Those quantities are denoted by a, b, c etc. these are elements of some field. That field is generally the field of real numbers R or the field of complex numbers C such that a (psi) is also an element of V. so you define multiplication by a scalar and it still gives you a vector. Just like in an ordinary three dimension space, I have a position vector r, twice r or thrice r exactly remains the same. This has all the obvious properties namely, a times b (psi) is the say as (ab) times psi.

(Refer Slide Time: 00:13:34 min)
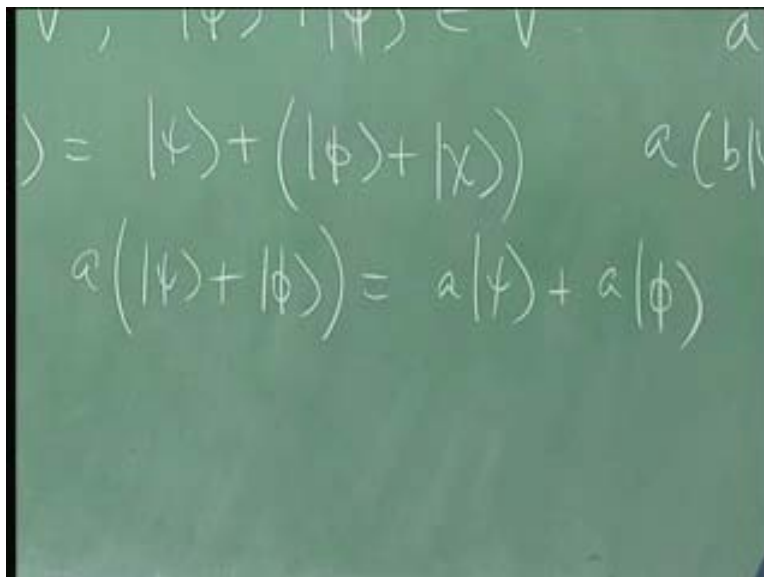
$$a|\psi\rangle \in V$$

$$a(b|\psi\rangle) = (ab)|\psi\rangle$$

We are not going to worry at the moment about not commutative fields. So we are going to assume its real or complex which is R or C. So that means that ab is same as ba.
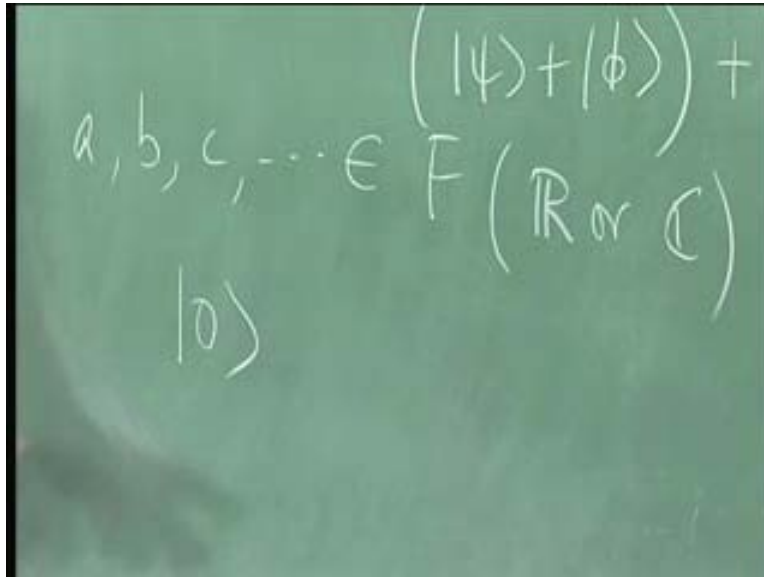
(Refer Slide Time: 00:14:29 min)

$$\rangle = |\psi\rangle + (|\phi\rangle + |\chi\rangle)$$

$$a(|\psi\rangle + |\phi\rangle) = a|\psi\rangle + a|\phi\rangle$$

Then the further properties are a times (psi + phi) = a (psi) + a (phi). So all the obvious properties, that you know about multiplying ordinary three dimensional vectors by real numbers are listed one after the other. There exists in this space a special vector called a null vector.
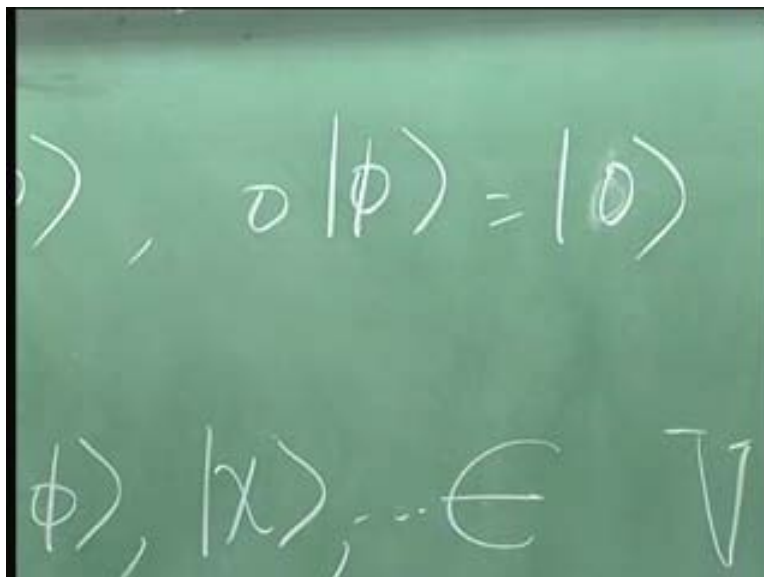
We are going to use this for something else but right now when we are taking about what a null vector is, let me just write it as ket 0. The reason I am hesitant to write it in this form is obvious. As you already know, when we solve some problems in quantum mechanics, we will be looking at various states and one of them is called the ground state.

It could correspond to the lowest energy for this system and occasional I will use this symbol 0 inside the ket for the ground state but that's not the null vector. This null vector is such that any vector phi + the null vector is still equal to phi.

Among scalars, there are two special scalars. There is a scalar where one times phi is of course phi itself and zero times phi is a null vector. A null vector is a vector where all of whose components are zero. The position vector at the orgin for example, is a null vector in three dimensional spaces. So since 0 times any vector phi is a null vector sometimes, what one does is, forgets about writing the null vector. One forgets about writing it as ket 0 and just uses ordinary 0 for the null vector because of this relationship. So you might as well just use ordinary 0 for the null vector but you shouldn't get confused because if I add two vectors, I still get a vector.

On the other hand, if I write (vector phi + vector psi) =0, what I mean on the right hand side is a null vector and not the scalar zero. But there should be no confusion because it is obvious from this fact that I could as well just call this zero. So these are all the properties that you need and since a runs over all reals for example, you can add the idea of multiplying vector by -1 and then get - the vector.

Now this set of properties defines a linear vector space and there are many examples of linear vector spaces all around. They have nothing to do with ordinary vectors and three dimensional spaces. These are abstract properties. For instance, the set of real numbers itself is a linear vector space. It is clear that these scalars would be the real numbers themselves and the vector would also be the real numbers. When you add two real numbers, you get another real number and you have a null vector which is this number 0 itself.
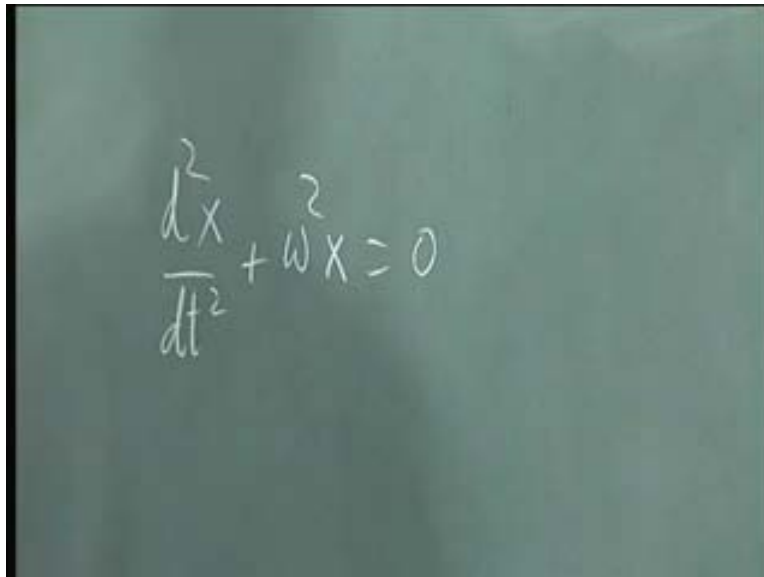
(Refer Slide Time: 00:18:32 min)
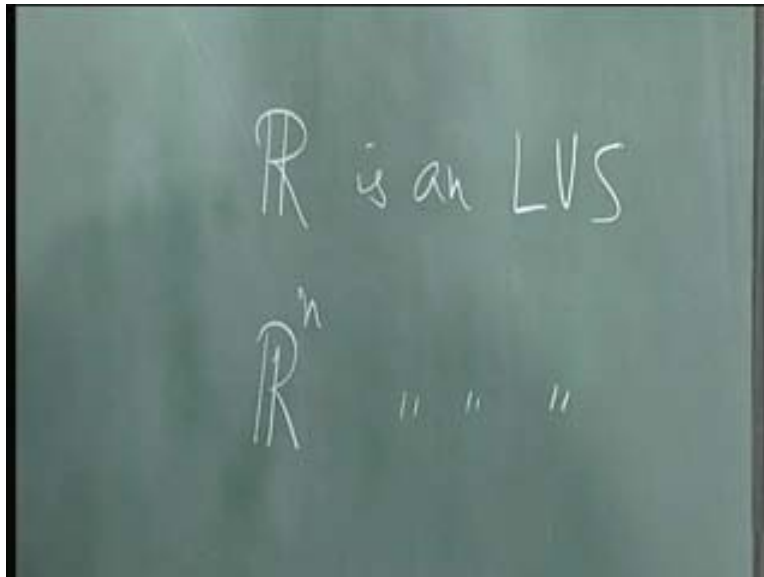
R is a linear vector space. I will call this LVS. The set of real number itself is a liner vector space. R2, points on a plane is a linear vector space so Rn is also a linear vector space. In other words, the set of x1, x2, x3 up to xn such that I define addition by saying

if you add two vectors with components x1, x2, x3 up to xn and y1, y2, up to yn, then the result vector is x1 + y1, x2 + y2 and so on. You add component wise. That itself forms a liner vector space. These are real linear vector spaces. When I go over to complex vector spaces, I would permit multiplication by complex numbers, for instance. Those are other properties that are used. They would form a complex linear vector space. These are the only two kinds of linear vector spaces we are going to look at.

There are other instances which are not so trivial. For instance, the set of all n by m matrices forms a linear vector space because when you add matrices, you get another matrix. When you add things to a null matrix, nothing happens to it. You multiply the matrices by scalar numbers; real or complex, you are still in that n by m matrix space. So the set of n by m matrices is a linear vector space.

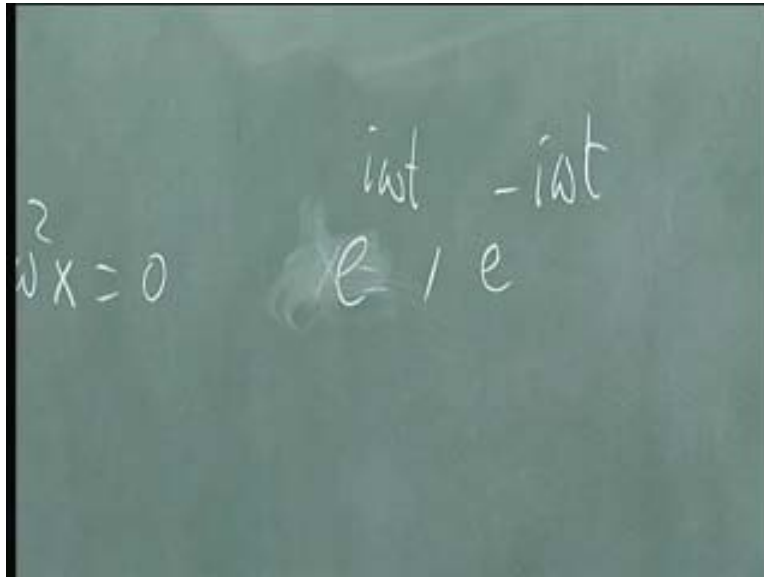(Refer Slide Time: 00:20:14 min)

There are other even more non-trivial examples of linear vector spaces. For instance, the set of solutions of the equation d squared x over dt squared + omega squared x =0. The set of solutions of this equation form a linear vector space because it's a linear equation. There are primitive solutions to this.

(Refer Slide Time: 00:20:27 min)

The solutions are e to the i omega t, e to -i omega t. they are the linearly independent solutions to this equation. You can form linear combinations of them and they still continue to be solutions. A very popular linear combination of this is called the cosine of omega t. Another is called the sin of omega t. but any a sin omega t + b cos omega t is also a solution and it satisfies all those axioms. So you see linear vector spaces can be quite general. The elements could be numbers, vectors in the ordinary space, vectors in n dimension Euclidean space, set of matrices and set of solutions of some differential equation and so on. So the concept is very general and extremely useful. So far, we have not introduced the idea of a distance between two vectors. We haven't introduced idea of the product of two vectors like a dot product. Those are all add-ons which come later but a linear vector doesn't need any of those. A very popular a way of representing ah elements of Rn is to write them in the form x_1, x_2 and to x_n in ordered n- tuple of real numbers.
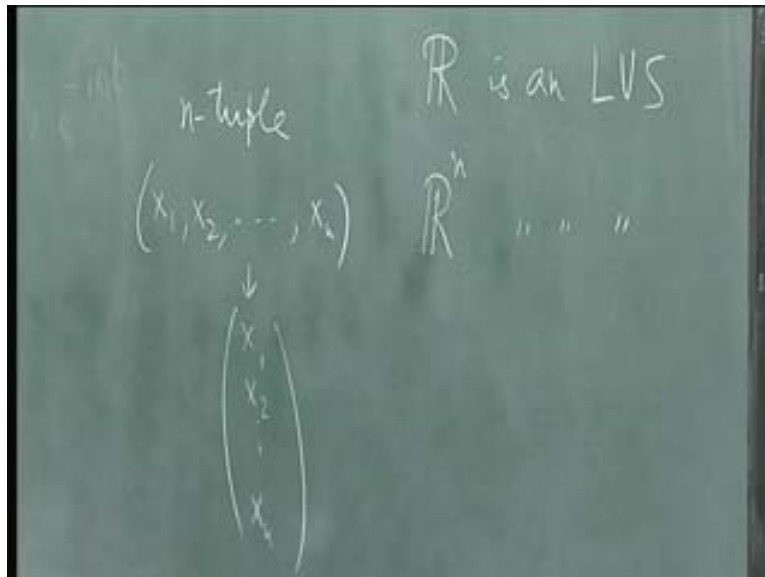
Now to perform manipulations with this, it's useful to represent this quantity in the following way. Write it as a column vector. That's one more way of representing an element of Rn. Now since you already know matrices, the moment you have a column vector, one is tempted to form a row vector and then multiply it on the left hand side to form a scalar. So the idea of the scalar product of two vectors emerges once you start putting this extra structure in. But it's not such a straight forward matter. So let's see how to go about that. Notice in particular I said we have not talked at all about the product of two vectors or the distance.

Just as Rn is a linear vector space, it's just the set of n-tuples of real numbers. We haven't said here is one element, there is another element and what's the distance between them. We haven't defined this quantity of distance. Once you put a matrix, then this Rn becomes Euclidean space here. So the first thing we want to do is to try so see whether we could introduce the concept of scalar product among these vectors. To do that, you need to recognize the following.

(Refer Slide Time: 00:24:44 min)
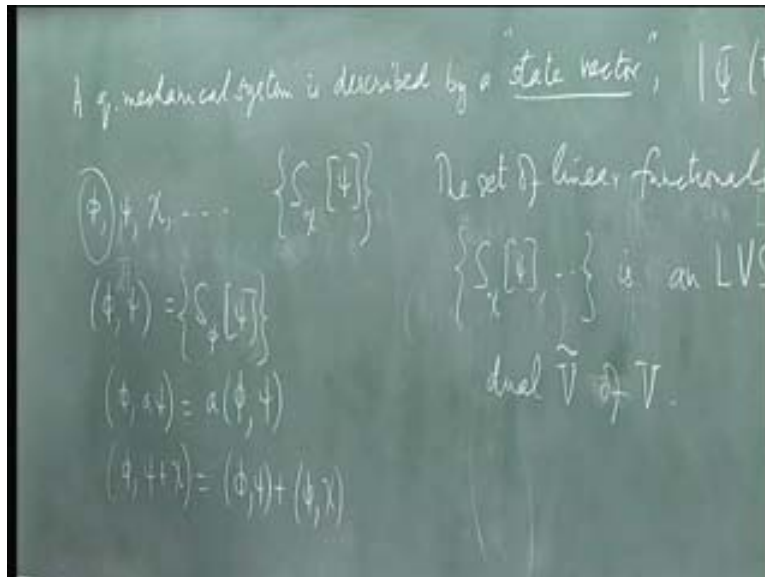
Let me for the moment drop this ket vector notation and write phi, psi, etc as it is, as the elements of the linear vector space for simplicity. Then I would like to associate with each vector a scalar. So I chose a particular special member of this linear vector space, let's say phi.

Let's take a special element phi and associate with each of the vectors in this space, a scalar which I denote by (phi, psi). This is equal to some scalar which depends on that special element phi and it is dependent on what's out there. I associate a scalar but I don't tell you at the moment how I associate a scalar. We will come to that. For every element of this vector space, I have chosen a special element phi and when I do something with that element and phi, I get a scalar number with a certain number of properties. The properties are that (phi, a psi) is the same as a times (phi, psi). (Phi, psi plus chi) equal to (phi, psi) plus (phi, chi).

So it says when you are doing this associating, if you take the sum of two vectors and you find the associated scalar, you might as well find the associated scalar for each of the members of the sum and add them up. By postulate, I say that this is equal to that (Refer Slide Time: 26:49) and similarly if you multiply this psi by some scalar number, you might as well have found this scalar and then multiplied it outside. So the moment you put in these properties, then for every element here you can find such a scalar. The set of these scalars over all the phi's here in the space form a linear vector space. The scalar here is not a function of psi, in the sense of f(x) or something. It depends on psi and the reference element phi. Now one can show that with these postulates, this set of quantities form a linear vector space itself. Well a function I want to reserve for things where I can differentiate integrate you know continuity and so on I don't want to use this is instead depends on this I have set of discrete elements So it's clear that it's discrete set of

elements and not a continuous function. We will define function more carefully later on. And it's a linear function. The whole idea is linearity because of this prosperity. It's a linear function. It doesn't depend on some squaring this or cubing this or taking logs or anything. And these forms themselves from a linear vector space and now you could say why I should choose S phi. I choose another element and I compute the set of S chi and so on. I choose all these chi's and all these put together form a linear vector space called the dual of the linear vector space.

(Refer Time Slide: 29:38)
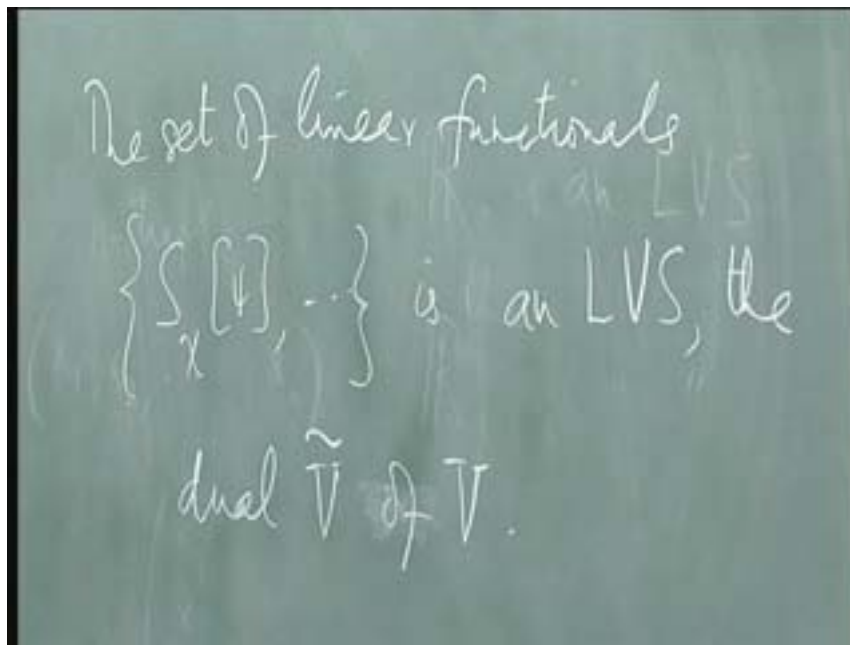
So the set of linear functions is an LVS and it's called the dual of the linear vector space.

In other words given a linear vector space, there is a natural mathematical way in which I can associate another linear vector space. So linear vector spaces come in pairs the structures these axioms are seen to it that there come in pairs of this kind.

Professor - I haven't told you the rule at all. I am just saying if you can associate with each psi a scalar by this specified rule, then that forms a linear vector space. As far as the existence of this dual is concerned, we don't need to know what the actual rule they have to satisfy is. Student- How do you reconstruct V?

Professor- A very good question! Can I reconstruct V from this dual? We will answer this. We will see by looking at examples. So it's a collection we always talk about let me write down what this is for Rn and then immediately you will see what this quantity is. There is unique dual to ever linear vector space I am not proving all these theorems. My idea is to show you the operational method by what I going to do Rn for instances. So these are theorems which exist.

Now the question is what the use of this is. There is a natural way to associate a scalar with a vector. as you know in ordinary three dimensional vector spaces, if you're given vectors like a, b, c etc and asked to construct a scalar from these vector, you do something called the dot product. So this quantity here (Refer Slide Time: 33:39) is the dot product. So we are heading towards a dot product of two vectors by associating a scalar with a pair of vectors. So it's a bilinear operation. We are taking 2 vectors and doing something to it to get a scalar. Now since S $_{chi}$, etc themselves form a linear vector space, it is convenient to say that this vector space here (Refer Time Slide: 34:14) which consists of a set of elements here is represented by writing this writing this psi in this form. This is an element of V and by saying that this phi here on the left hand side in this bracket, let me write it in this form and say that this is also a different kind of vector which lives in a dual space.

(Refer Slide Time: 34:46)
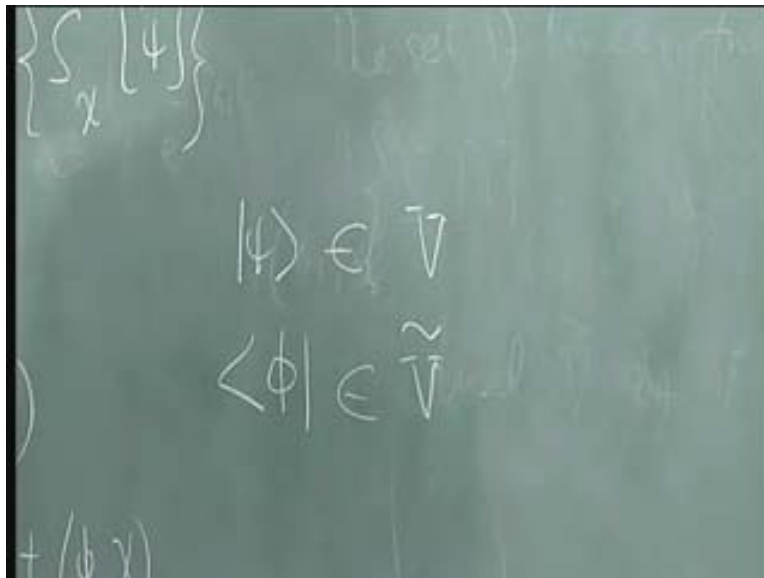
In other words, I have a set of scalars and now I am saying for every for every number which is in this space (Refer Slide Time: 35:03) they are created by taking the original ket vectors elements of V, doing something to that using this reference element phi. So this set of numbers I now write in a given notation instead of this kind of bracket and I say this is an element of dual space and are called bra vectors. These are called ket vectors (Refer Slide Time: 36:10).

(Refer Slide Time: 36:10)
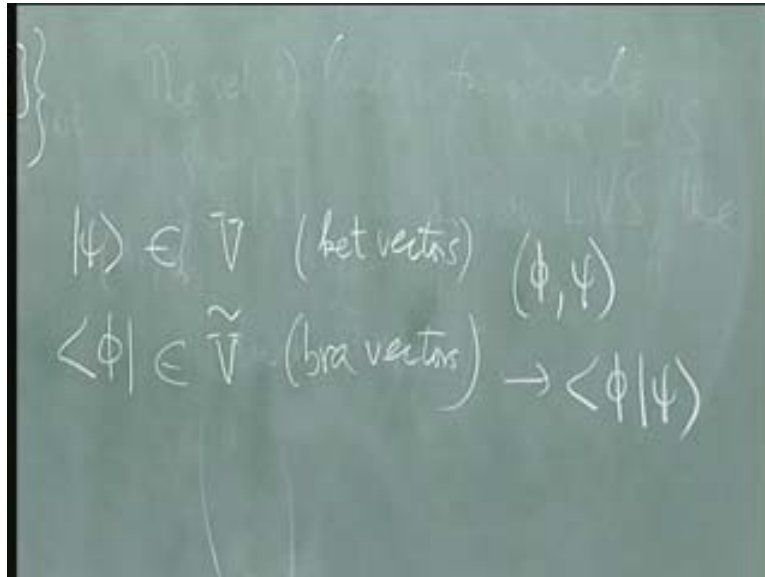
This is the notation used in physics. So what you have done is to take the set of linear functions here and replace them by an element of a dual space phi and then you say this quantity here stands for the linear function or the scalar. So that dual vectors are going to be written in a different form from the ket vectors from V, so that I know that these belongs to V. This involves phi they can involve many other vectors all those vectors all these guys I now say a represented by this phi here um with chi phi for every element every element in the original vector space there is an corresponding element in the dual vector space here such that I take one element from the dual space I take one element from the original space and I form a bilinear combination to give me a scale which satisfies these properties.
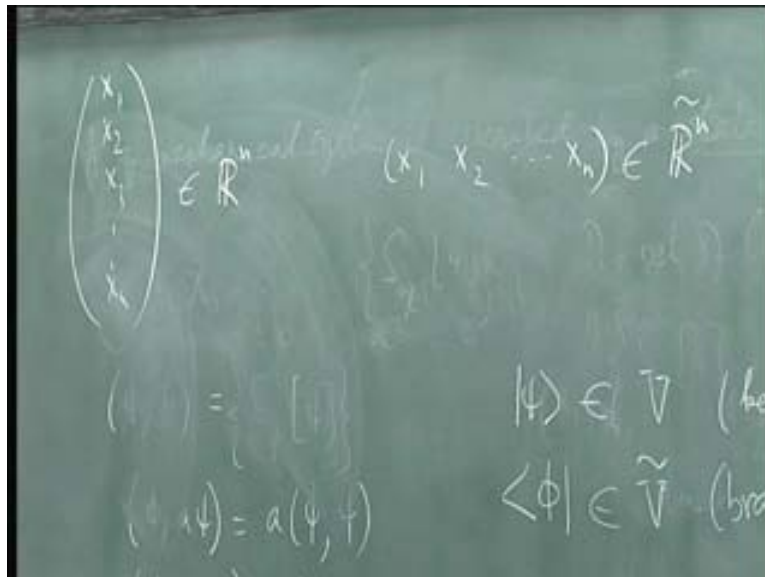
I write $x_1, x_2, x_3\ldots x_n$ are elements of Rn. if I construct the corresponding row vector in this form, then I say all the vectors are elements of the dual to Rn. Then there is a natural way to produce a scalar which satisfies these properties which is matrix multiplication with the bra vector on the left and ket vector on the right or the row vector on the left the column vector on the right.
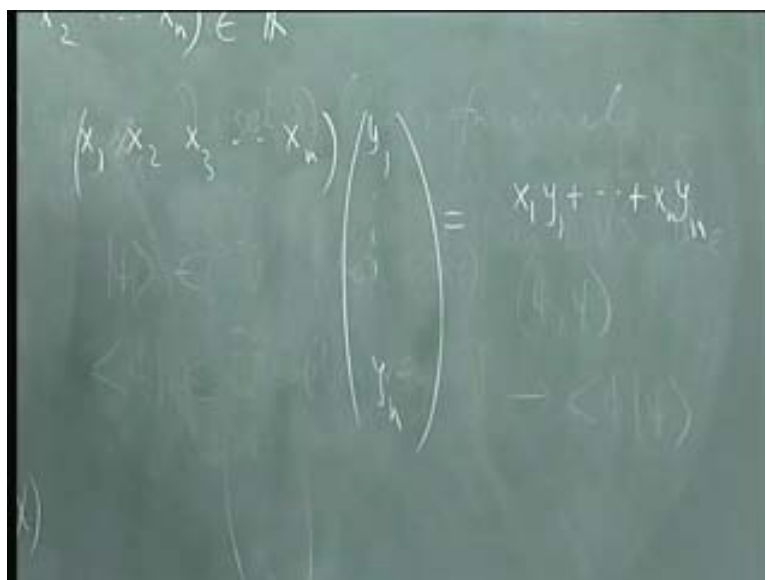
So it's immediately clear that $x_1, x_2, x_3 \ldots x_n$ with say y_1 to y_n gives a scalar. its $x_1 y_1 + x_2 y_2 +$ etc up to $x_n y_n$. It's intuitively clear that the dimension of the dual space must be the same as the dimension of the original space. Since I have said that it's in one to one correspondence, for every element in the original vector space, there is a corresponding element in the dual vector space. And the way to remember it is to say that one of them is represented by column vectors the other by row vectors.

I could have changed this notation. I could have said the original space is represented by row vectors and the other by column vectors. It would make multiplication a nuisance. You have to be careful how you multiply. Since I multiply row by column and that's the rule by matrices. So we will use this convention. The further complication which looks a little confusing is that this space here is exactly the same as the original space itself. It so happens that Rn is the same as its dual. Still we will like to distinguish between the space and its dual and therefore I will represent one of them by column vectors by and the other by row vectors.

In fact there is an exact theorem which says every n dimensional linear vector space is isomorphic to n dimensional Euclidean space. In other words you can think of every n dimensional vector space in terms of column vectors and the corresponding row vectors. Its only when you go to infinite dimensional spaces that you run into some technique which are non trivial but all finite dimensional spaces look exactly the same. You cannot define the multiplication of two vectors belonging to the same vector space. So every time you take a dot product of two vectors in ordinary three dimensional Euclidean spaces you are really taking one element from the dual space one element from the original vector space and taking the dot product. As you can see you cannot multiply two column vectors and get a scalar. To get a scalar you need to have a row on the left and column on the right. It's just that in ordinary three dimension space these two are the same spaces.

So one doesn't realize that one is doing this but if you have to write it out in terms of column vectors and row vectors, its quite clear that you form a scalar or take a dot b. The way you write a dot b is $a_1 b_1 + a_2 b_2 + a_3 b_3$. If you represent a and b by column vectors, the a on the left hand side it has to be a row vector.

So the scalar product between these vectors is only defined by taking one element from V tilda and one element from V. It will become particularly important when you look at infinite dimension spaces then you have to very careful that you just that. You can define the product of a vector with itself but it's a map to something else.
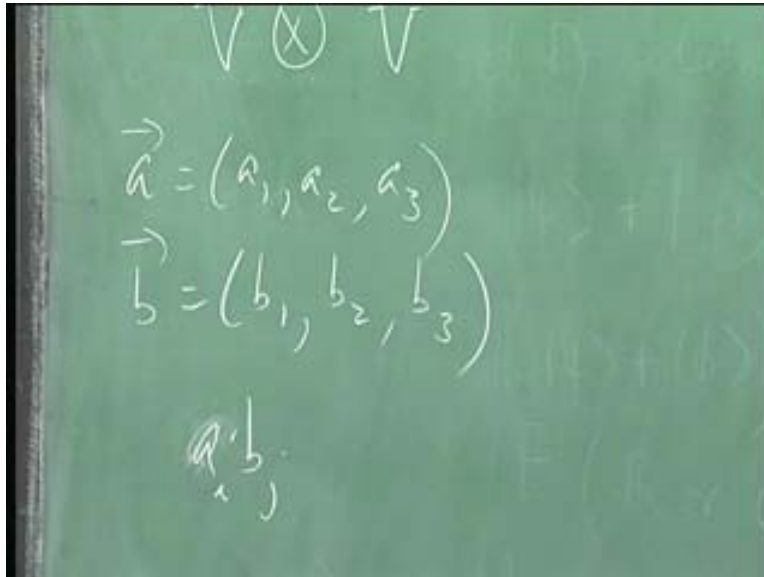
You can define a V direct product with V. So I would take a given vector and I take another vector from the same space and this would be mapped from the original vector space. That's a different space all together. Its dimensionality is higher. If each of this is n dimensional, then the dimensionality of the other space is n squared.

if you take ordinary Cartesian vectors with components like a,b,c etc and I write components, the set of numbers $a_i b_j$ is precisely this(Refer Slide Time: 46:20). It's an element of V cross V because if I take a_1, a_2, a_3 and I consider this set of numbers, $a_i b_j$ have nine possibilities. So immediately you see that this quantity ai bj is not an element of the original R_3.its an element of $(R_3$ cross $R_3)$. It's a nine dimensional space. I write it like a vector b vector without dot or a cross. They are called Cartesian products or tensor products. But we are taking about finding a scalar from these vectors and for a scalar, you take an element of the dual, you take an element of the original space and you multiply. Now the moment I do this, I also have the possibility of defining the inner product.

(Refer Slide Time: 00:47:27 min)
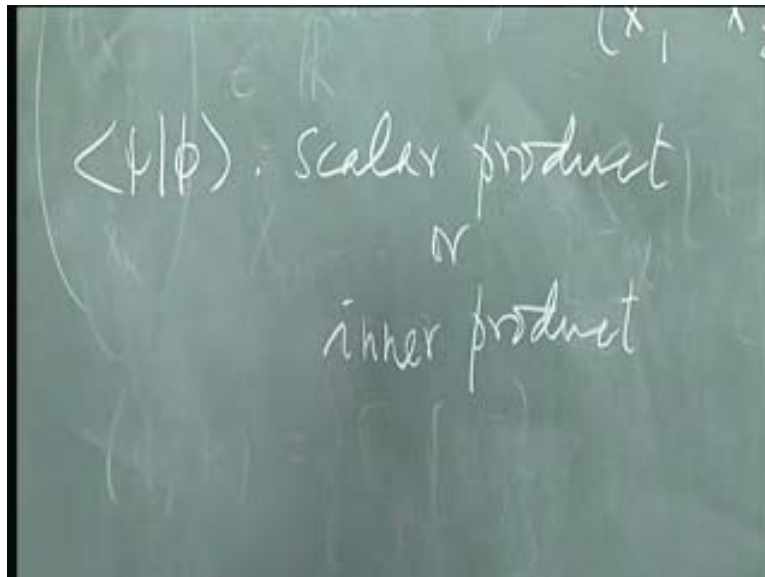
It has the properties we wrote down for S. the moment I do this, I can define the inner product of a vector with itself. Let us the take the inner product of psi with itself. What would this be in order in three dimensional euclidean space or n dimensional euclidean space?

(Refer Slide Time: 00:48:20 min)
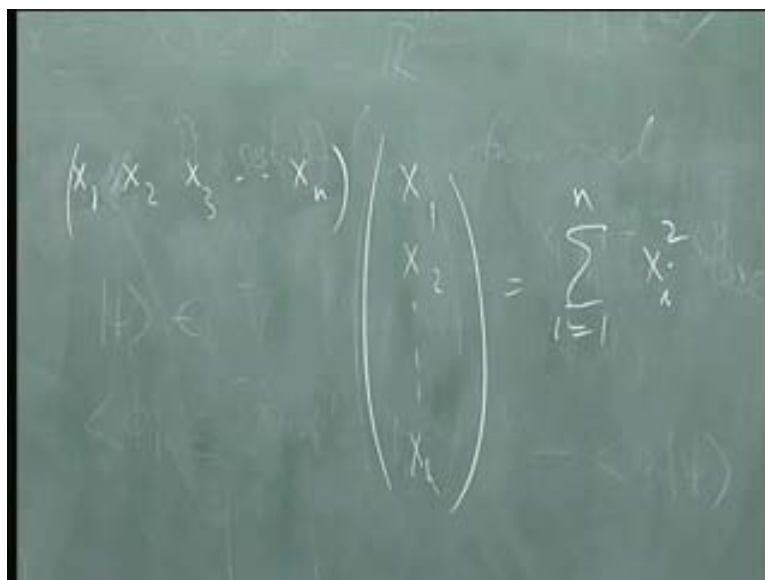
If psi is represented by x_1, x_2, up to x_n, then this quantity here is equal to summations I equal to 1 to n, X_i squared. This corresponds to the length of this vector. So if I say it's a

vector in n dimensional space with component x_1 to x_n then sum of the squares of the element is the square of the length of this vector and because it is a sum of real number of positive quantities, it would be zeros if and only if the vector is a null vector. So this is equal to 0 if and only if psi is the null vector. We would like to preserve this property but then we said that these vectors are completely general quanties and could in fact be multiplied by complex numbers. Then this is no longer true. How would you preserve that? I define the element in the dual space by taking the complex conjugate.

(Refer Slide Time: 00:49:55 min)

So if these(Refer Slide Time: 49:55) are not elements of Rn but are elements of a general n dimensional vector space, to avoid confusion let me use some other symbol for it. Let's call it alpha$_1$ to alpha $_n$. If this corresponds to psi, this psi should really be alpha$_1$*, alpha$_2$ * to alpha $_n$ *. Now they are in good shape because this corresponds to mod phi the whole squared. And you are guaranteed that it is zero if and only if each of the alphas is zero. so we begin to see our first generalization that if you are looking at complex vector space and you have an n dimensional space, the elements are in a column vector in the dual space which are the bra vectors or the complex conjugate transposes. So this is the reason why in matrix analysis complex conjugation is not a very natural operation. Complex conjugate transposes is a natural operation and it's called the Hermitian conjugate when you take complex conjugate transpose.

(Refer Slide Time: 51:50)

This relation is still true but it also implies that if you have a vector a on psi and this is equal to psi prime say, then for the corresponding psi prime, you have to take each element of psi, write it as a column vector, multiply by a and take complex conjugate transpose. So it is immediately clear that is equal to a star on psi. So any scalar you multiply, if its multiplying an element of the dual space, you start with an original vector multiplied by a scalar.

You have a new vector. If you first want to find its adjoint in the dual space, then take the adjoint of the original vector and multiply by the complex conjugate of the number of the scalar and this will satisfy all those associative properties and so on. The moment of you have this, you can now start defining the distance between two vectors because you have the idea of a scalar product. Notice that this rule also implies that if I take the complex conjugate of this scalar phi with psi, this is equal to psi phi. So scalar products in general don't have to be real. It could be complex numbers but the inner product of phi with psi is not the same as the inner product of psi with phi because there is a complex conjugation involved here.

(Refer Slide Time: 00:52:47 min)

So that's missing in real vector spaces and that's why you write a dot b = b dot a. that's not true in general. a dot b is b* dot a. but in a real vector space, we are multiplying only real numbers. So this complex conjugation doesn't make a difference. Now I would like to define the norm of a vector and I denote it like this (Refer Slide Time: 54:26). This is by definition, half the positive square root of this positive number psi with psi. This is a non negative number and it's equal to 0 if and only if psi is a null vector.

(Refer Slide Time: 00:54:04 min)

Now we are going through things which are exactly the same as what happens in ordinary three dimensional Euclidean space and those concepts should remain when you generalize this system. Therefore one would like to have a statement like psi + phi which is the norm of the length of the sum of the two vectors. This obeys a triangle inequality. We know in ordinary three dimension space, the sum of two sides of a triangle is greater than the third always unless a triangle collapses. So this quantity is lesser than or equal to psi plus phi.

(Refer Slide Time: 00:56:30 min)

$$\vec{a} \cdot \vec{b} = a\,b\,\cos\theta$$

$$|\vec{a} \cdot \vec{b}| \le |\vec{a}|\,|\vec{b}|$$

You also know another thing that if you took ordinary vectors a vector I dotted it with b, in the usual real vector space this quantity is a b cos theta by definition. Theta is the angle between them and the magnitude of cos theta is between minus 1 and plus 1. It's equal to plus 1 if the angle is 0 and minus 1 if the angle is pi. That's the extent to which it can vary. Therefore it follows that the norm of the magnitude of a dot b is less then or equal to the magnitude of a times magnitude of b. that's just a statement that the cosine has a valve between minus 1 and minus 1. Now that's generalized to this and the statement is that phi with psi mod square is less then equal to the norm of phi with phi psi with psi. Since this is a complex number in general, we need mod squared. This has a name and is called the Cauchy-Schwarz Inequality. We use it very extensively to establish the uncertainty principle. At the mathematical level it follows from the Cauchy Schwarz inequality. Now you could ask when is this inequality an equality. This is only when a and b are in the same direction or antiparallel. In other words when, they are collinear.

(Refer Slide Time: 00:58:12 min)
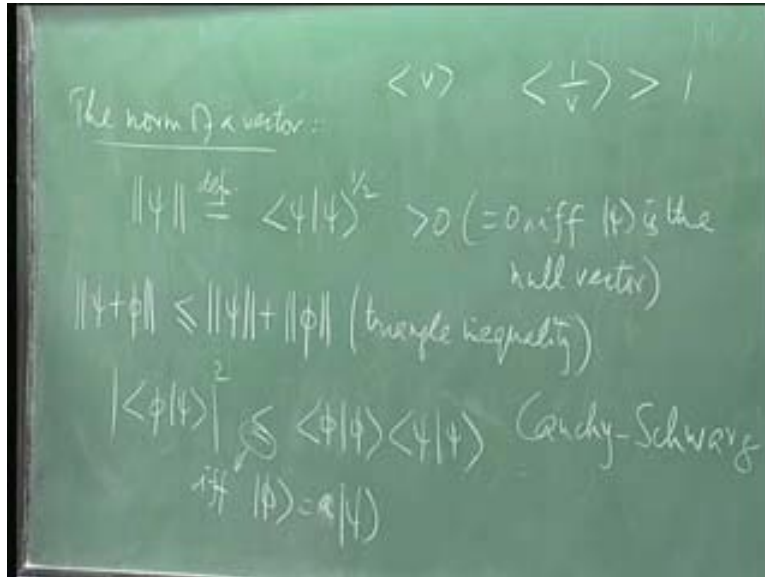
The same thing is true here and what does it mean to say two vectors are collinear. It means one of them is a scalar multiple of the other. The direction is the same. In exactly the same way, this thing becomes an equality if and only if phi for instance, the ket vector phi is in the same direction as the ket vector psi. In other words, its just psi multiplied by a number. What does it mean by linearly dependent? Psi itself is said to be made up of adding several vectors. Every vector can be decomposed to other vectors. So phi is linearly dependent on psi. Then the Cauchy Schwarz inequality becomes an equality. Otherwise it remains strictly less than this quantity. We will see how powerful this statement is. Just to give an example from a way out from this whole thing. You take the gas in this room. it obeys a Maxwellian distribution of velocities and you can compute the average speed. So lets compute for instance, the average speed and let's called it v. if I compute this quantity v, this depends on the square root of temperature. Its some root kt over m where m is a molecular mass, k is Boltzmann constant multiplied by some number.

You could also ask what about 1 over v. what about the average value of the reciprocal of this velocity the of the speed? It's clear that 1 over the average is not the same as the average of 1 over the speed and in this case, it is strictly greater than 1. This can be shown very trivially by using the Cauchy-Schwarz inequality. It's a one line proof and we will do that at some stage. So just to show you this inequality which starts off very innocuously with just the scalar product of two vectors in ordinary space as profound implication its part of a much deeper fact. You can generalize this to a set of n vectors at a time and arbitrary n number. That brings us to the concept of linear dependence which will then bring us to the concept of basis set in this vectors space, expansions basis sets, orthogonalization so on. We will talk about that next time. Thank you!