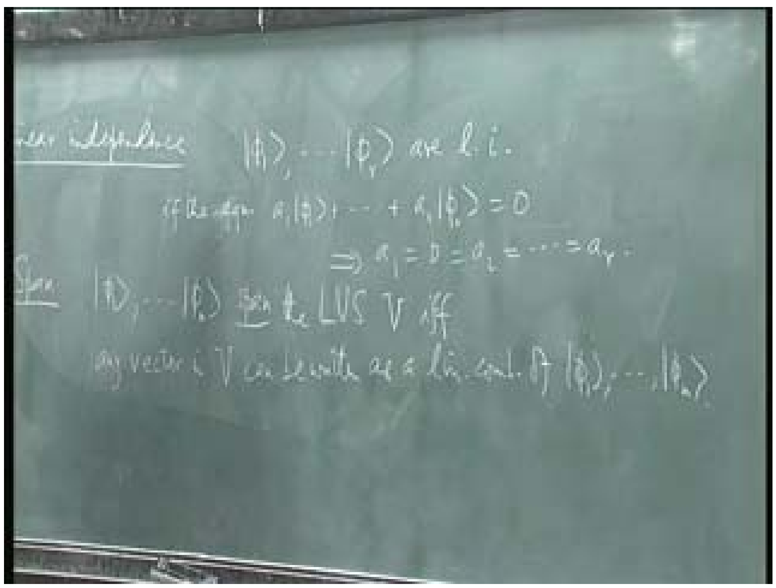
Quantum Physics

Prof. V. Balakrishnan
Department of Physics
Indian Institute of Technology, Madras

Lecture No. # 03

We now continue with our preliminaries regarding linear vector space. I would like to introduce the idea today of basis set in a linear vector space. For that, you need two concepts. One is that of linear independence.

(Refer Slide Time: 00:01:24 min)

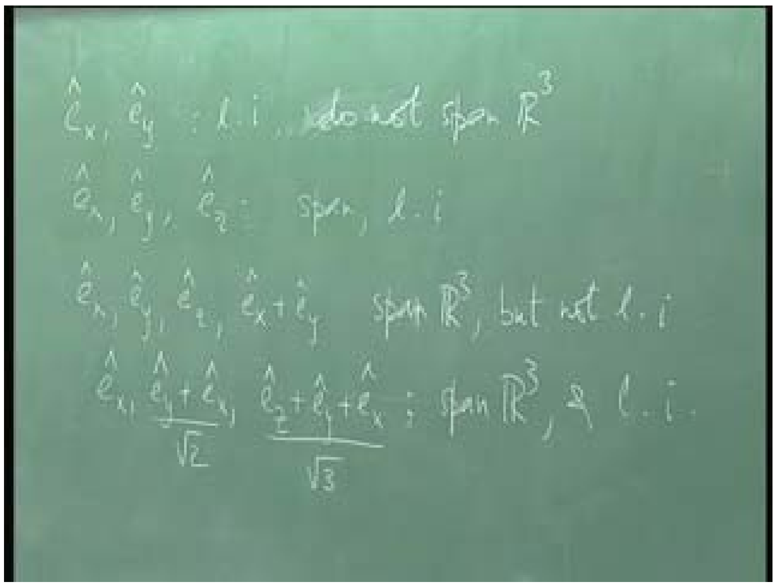


I say a vector is linearly independent of another set of vectors if the former cannot be written as a linear combination of the later. so phi_1, phi_r are linearly independent if the equation a_1 phi_1+ etc… + a_r phi_r =0. This equation implies that a_1= 0= a_2 = a_r. The only way we can satisfy this linear equation is we 0 the coefficients. Then you say that this set of vectors is linearly independent. This is a crucial concept in linear vector spaces that a linear independence. If such an equation has a solution with non-zero coefficients, then you say there are linear relations between these vectors and some of them are linearly dependent on the others.

In particular, if you have just two vectors, a_1 phi_1 + a_2 phi_2 = 0, it implies that phi_1 is $-a_2$ over a_1 times phi_2. In another words, in geometrical language you would say the two vectors are along the same direction. So that's the first concept. The second thing you need in order to define concepts such as dimensionality and so on is that of the span of a set of vectors. A set of vectors in a linear vector space is said to span the space if every vector can be written as a linear combination of these vectors.

So again let me say phi $_{1 \ldots\ldots}$ phi $_n$ span the linear vector space V if and only if any vector in V can be written as a linear combination of these vectors. So this is the idea behind a set of vectors spanning this space. Any vector in that space should be writeable as a linear combination of these given vectors. The set of vectors spans this space. These are independent concepts. One should confuse one for the other. For example, in ordinary three dimensional Euclidian space, e_x, e_y are unit vectors in the x and y directions.

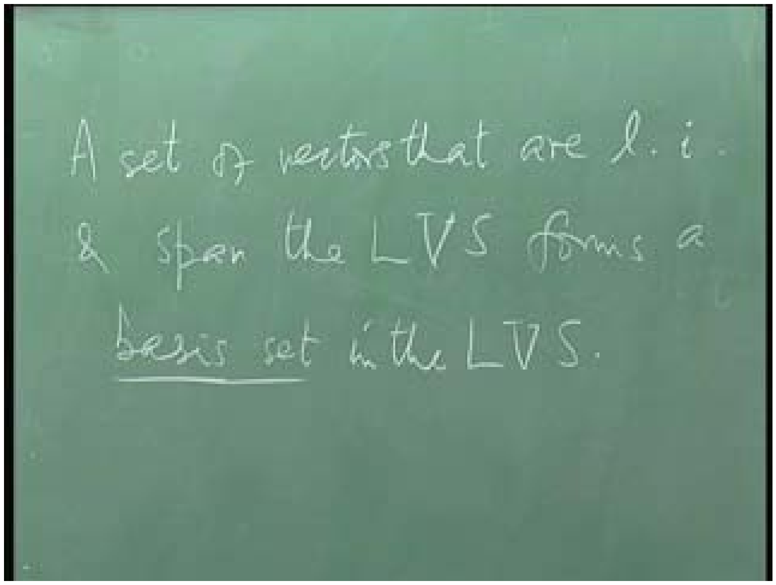
(Refer Slide Time: 00:04:44 min)



They are linearly independent of each other because e_x is certainly not a multiple of e_y but they do not span R3. e_x, e_y and e_z span the space and are also linearly independent. None of them can be written as a linear combination of the other two. Here is a set e_x, e_y, e_z, e_x+ e_y which spans the space R3 but are not linearly independent. How about this set e_x, e_y + e_x, e_z + e_y + e_x? Do they span the space? Yes they do and are they linearly independent? Yes they are linearly independent. What's the difference between this set and this set here? (Refer Slide Time: 07:18).they both span R3 and they are both linearly independent.

One set is orthogonal and the other set is not orthogonal. It's like choosing oblique coordinates in three dimensions. This is like saying I choose the x axis, this vector e_y+e_x here which is a 45degree line in the xy plane and e_x+e_y+e_z is at an angle to the xyz axis. I choose this sort of a set of axes. So it's possible to have the set of vectors which span the vector space and are linearly independent but are not orthogonal. There is another difference between these 2 sets of vectors. They are not normalized. The magnitude of e_y+e_x is not 1 or this vector e_z+e_y+e_x is not 1. You can normalize these vectors trivially. Vector e_y+$e_{z\ \text{over}}$ root 2 and vector e_z+e_y+e_x normalizes them. They are normalized to unity but they are not orthogonal. Now if there were both normal normalized to unity each of them and orthogonal then that would be like a Cartesian basis.

(Refer Slide Time: 00:09:04 min)

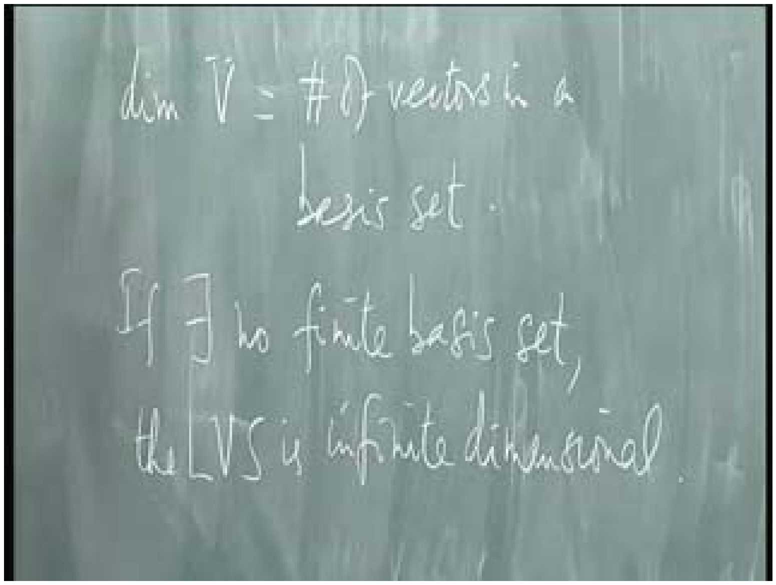


So a set of vectors that are linearly independent and span the LVS forms basis of the basis set. So whenever I say here the set of vectors form a basis in this a vector space, I mean a set of vectors that both spans the space and are also linearly independent each of other. So both are needed and I just showed you that these are not the same property. They could be mutually exclusive in some sets. They could be an example where one excludes the other.

So linear independence does not imply span and vice versa. You need both of them in order to form a basis. If in addition, the basis consists of mutually orthogonal vectors whose scalar products of different vectors is 0 and each of them has magnitude 1, then we call it an orthonormal basis. Orthonormal basis is where each vector has magnitude unity and different vectors from the basis set have zero scalar products. So I would write this as phi_i phi_j = $delta_{ij.}$

Now I go to abstract notation. These are the inner products. The inner products are zero if i is not equal j and equal to 1 if i is equal to j. So if the magnitude of the vector is unity then I say it's an orthonormal basis. I will very frequently use orthonormal basis sets but have other kinds of basis sets. We need to ask how many such vectors are needed. The number of vectors in the basis set in the linear vector space is called the dimensionality of this base and that's unique. For example, in three dimension Euclidean space you need three mutually linearly independent vectors which span the space in order to form a basis. That basis is not unique. This one is not an orthogonal basis. This one is but the number is three each case. So that gives you the concept of dimensionality of the linear vector space.

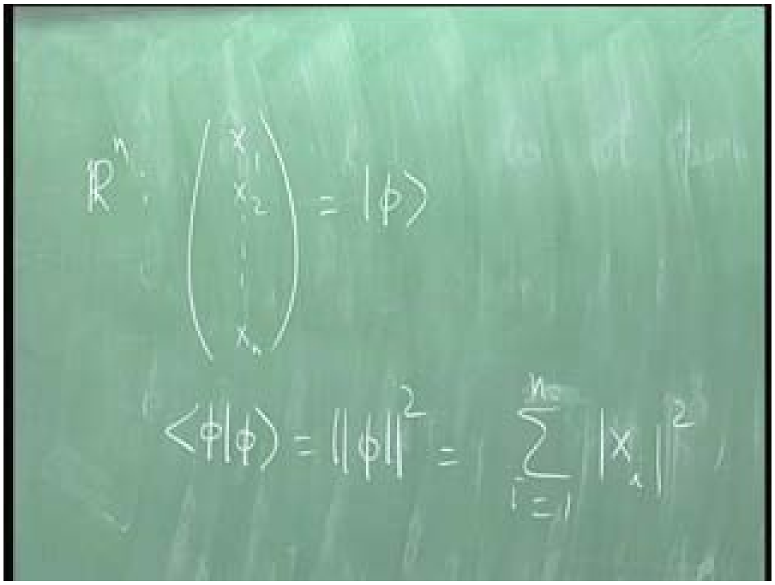
(Refer Slide Time: 00:12:13 min)



Given a non-orthogonal basis, we can always make it an orthogonal basis in construction which I will show. But the number is fixed for every linear vector space. If it turns out that you don't you have a finite basis, for any n there are vectors which are not in the span of the vectors which are already written down. Then you say this space has infinite dimensional. So an infinite dimensional space is one which doesn't have a finite basis set. Is R2 an infinite dimensional vector space? No when there is infinite number of vectors in R2 but every one of them can be written in terms of a linear combination of two non-collinear vectors. Rn is n dimensional and if you cannot find a finite basis set no matter how large n is, then you say this basis is infinite dimensional. It's like having a space with an infinite number of independent directions. If you look at R1e_x forms a basis because any vector in R_1 can be written as a number times of e_x, positive or negative.

That's a linear space but it has only one independent direction, the x axis. But we are not talking about an infinite dimensional space. Then you got to be a little careful. It's not obvious that all the things that you do for ordinary linear vector spaces would work in an infinite dimensional space. Here is what can go wrong. So let's look at Rn and try to make it R infinity by increasing the number of entries.

(Refer Slide Time: 00:16:05 min)

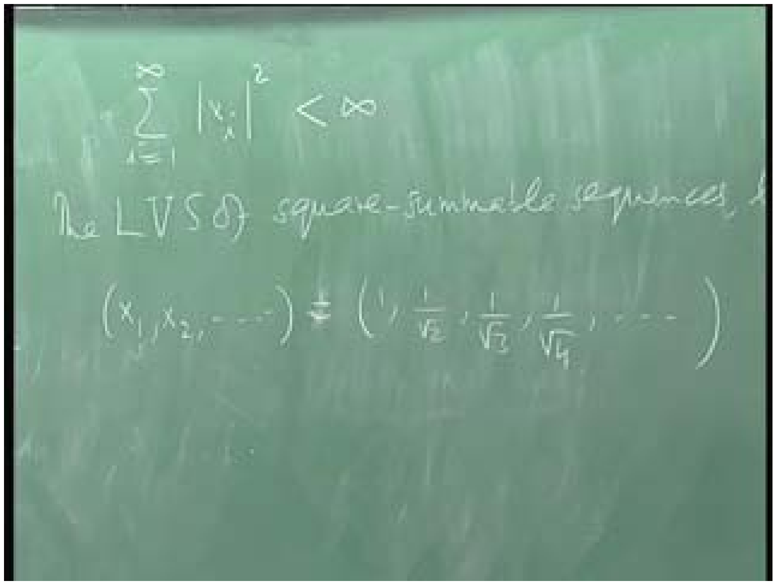


So any element of Rn you could write in the form x_1, x_2… x_n. This is equal to say phi.

Then phi with phi is equal to norm of the vector squared by definition and that is equal to a summation from i= 1 to n x_i mod squared. If the x's are in real vector space, then this is just x_i squared. This is sum of the square of the length of the vector in n dimensional Euclidean space. Now of course I would like to make this infinite but then there is not guarantee this converges. I would like have vectors of finite length.

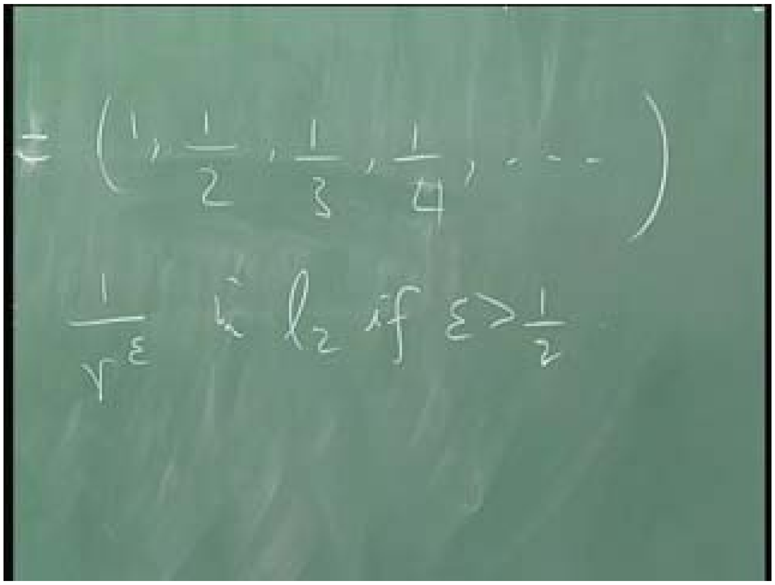
So this is not at all clear that this will converge when n goes to infinity. If it doesn't converge this doesn't make sense. When you add to infinities you get another infinity and so on. So I have to put in the condition that the vectors have finite length. To do that we would have to say that in R infinity, namely the space of sequences; x_1, x_2, x_3, up to x_n where n tends to infinity should be such that this quantity converges when n tends to infinity.

(Refer Slide Time: 00:17:40 min)



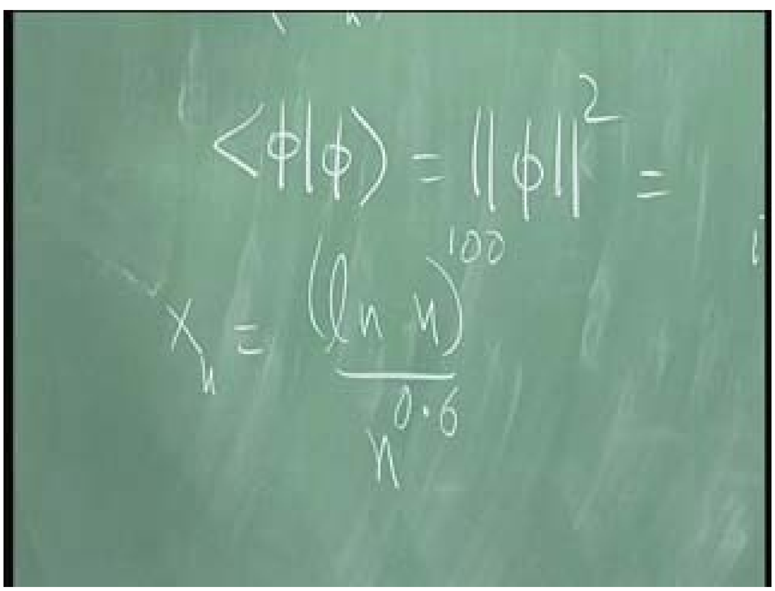
Then you require that you have a space in which i=1 to infinity x_i mod squared is less than infinity is finite. Then we can speak of respectable vector space of infinite dimensions with denumerable components x_1, x_2, x_3, x_4 all way to infinity but you require this to be finite. So it's a simple matter to prove that this is needed for the triangle inequality, the Cauchy-Schwarz condition to be true and so on. This space has a special name. This is the linear vector space of square-summable sequences and it's denoted by l_2. Then the triangle inequality is valid. Consider (x_1, x_2, etc) = say (1, 1/ root 2, 1/root 3, 1/root 4, etc). Is this an l_2? If I sum this, I square it first and sum it. 1+ 1/2+ 1/3+1/4 and so on is a harmonic series and this diverges. So it's infinite. So this is not an element of l_2.

(Refer Slide Time: 00:19:39 min)



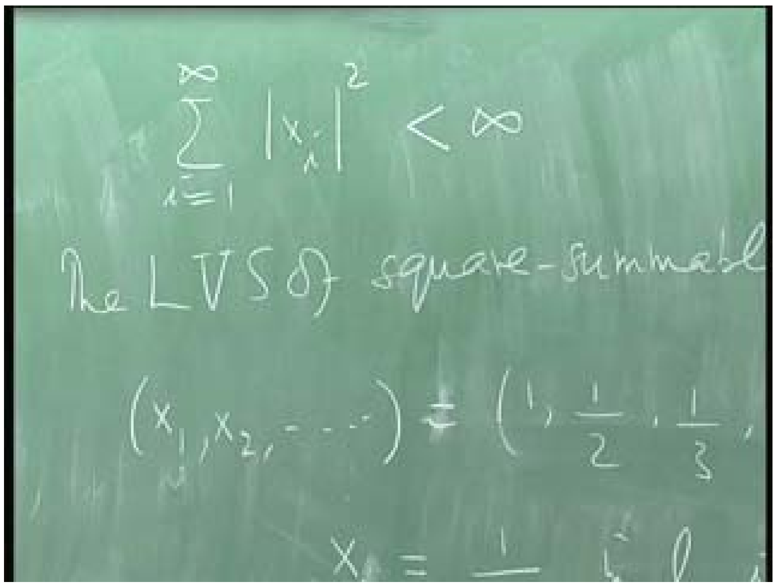
Now consider x_r= 1 over r power. This will be in l_2 only when epsilon is greater than half uh because when you square it you get 1 over 2 epsilon and that number 2 epsilon must be greater than 1 for it to converge. What happens if you have x_n = log n over n to the power 0.6? Is this going to be in l_2?

(Refer Slide Time: 00:20:52 min)



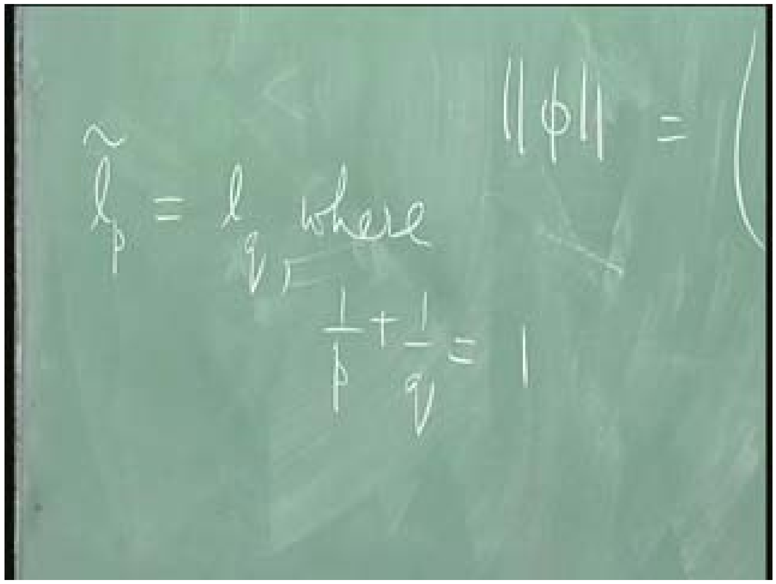
Remember in the denominator when square you get n to the power 1.2 which is greater than 1. So if you didn't have a numerator it would of course converge but you have log on top and log infinity of course is infinity. So is this convergent? Yeah it is. All of calculus can be summarized in one line. The log is weaker than a power and the power is weaker than an exponential. Certainly this converges and is very much in l_2. How about x_n = (log n) power 100 over n power 0.6? This will converge no matter what you raise it to. All you have to do is test this convergence of these quantities and then once it is valid, this is true.

(Refer Slide Time: 00:21:54 min)



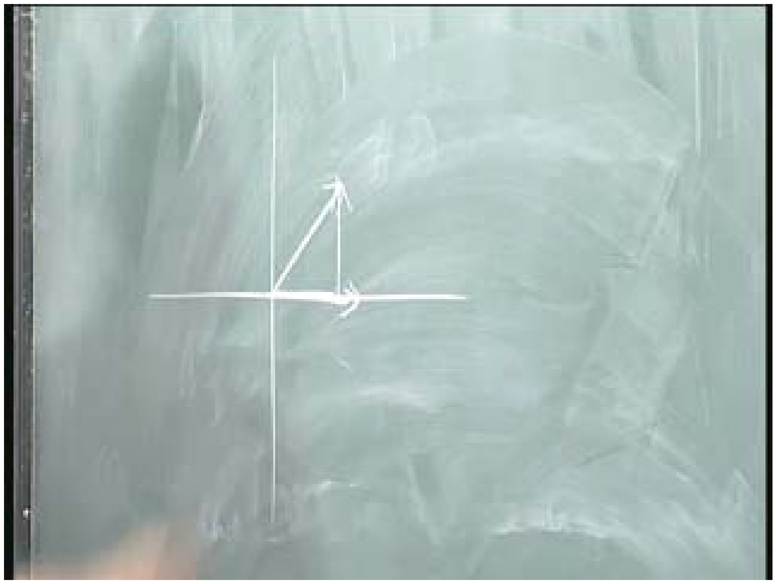
Now we are going to use square summable sequences. One of the reasons for doing so is that you can actually generalize this. You could ask why I have to do this. I could define norm in a slightly different way. I could for example, instead of writing the norm of a vector as equal to this (Refer Slide Time: 22:27) to the power half which is what the norm would be because a square of it is this. I could put a p here and put 1 over p where p is some positive number. I can define the norm of this kind. This would be l_p. I should probably write subscript 1 over p or something like that depending on the notation. The advantage of l two is that itself dual. The dual space is also the space of square summable sequences but the space l_p is not self dual.

(Refer Slide Time: 23:42 min)



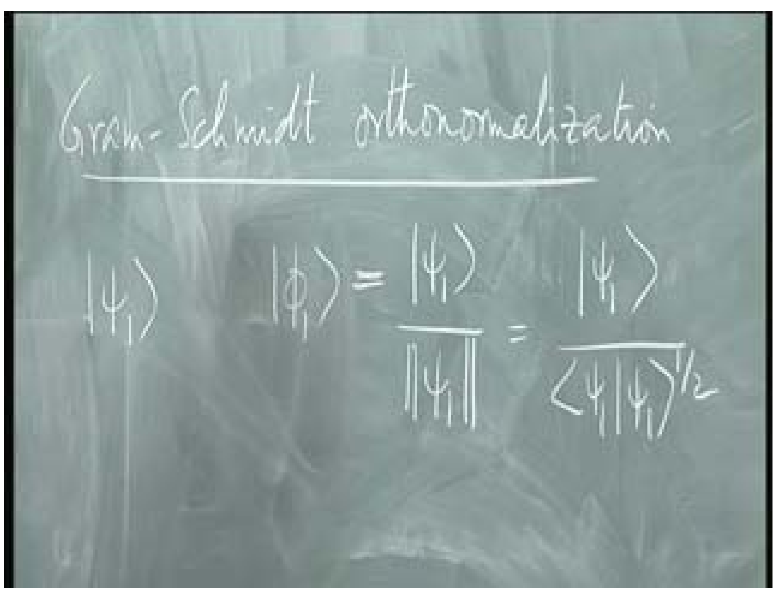
The dual to l_p is l_q where 1 over p + 1over q = 1. So if you improve in p, you go down in q and so on. So the only one which is self dual is when p=q=2. So we will restrict our attention to l_2, square-summable sequences. There is another physical reason for it because in quantum mechanics you are going to give a probability interpretation to the various inner products and so on and they naturally involve l_2 in the natural way. Hence we will not talk about p summable sequences. So we can define an infinite dimensional space in this form. Now let's ask what the advantage of writing a base is. First let me show you that if you start with an arbitrary basis which is not orthonormal you can make it orthonormal always. That's like starting with oblique coordinates in two dimensions, three dimensions, etc and then saying I am going to make it an orthogonal set.

(Refer Slide Time: 25:25 min)



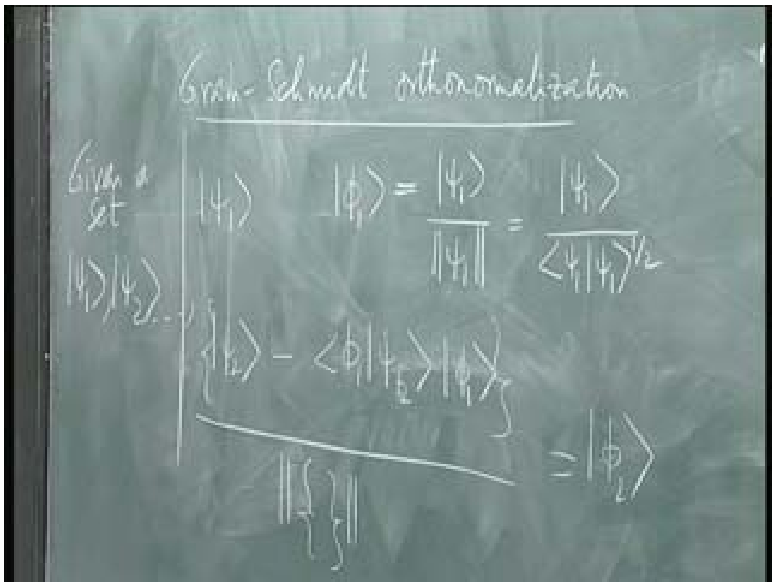
Now what you do in a plane for example, we are stuck with the oblique coordinates? the way you would make it orthogonal is to choose the first vector and call that unit vector in one direction and then take the second vector, project it down to this direction (Refer Slide Time: 25:22). And then I get rid of this part of it which is already included in the first direction and use this direction as the orthogonal vector and normalize it suitably. So this procedure done systematically starting with the 1st vector, 2nd, vector, etc is called Gram-Schmidt orthonormalization.

(Refer Slide Time: 00:25:40 min)



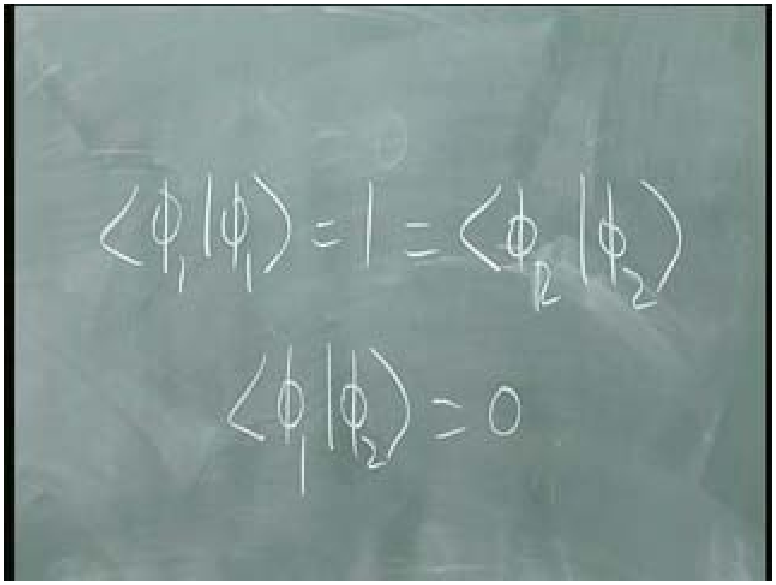
And it says suppose you are given with psi_1 to start with the first vector and then choose a phi_1 = psi_1 divided by the norm of psi one. Then remember that this is equal to psi_1 divided by psi_1 with $psi_1$to the power half. So you first you normalize it. This guarantees that the inner product of phi_1 with itself is unity.

(Refer Slide Time: 00:26:39 min)



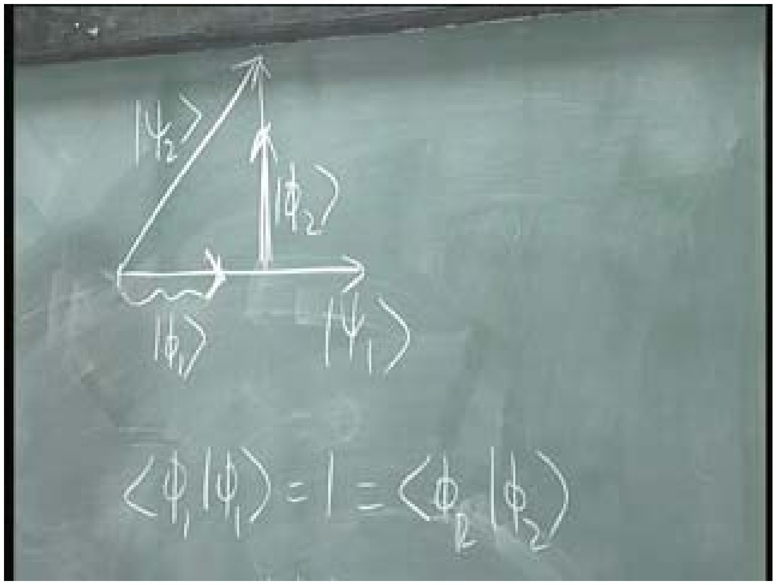
When you take the second vector psi_2 and you subtract from this psi_2 the portion of psi_2 that's along phi_1. This would be phi_1 with psi_2 which is a component and the vector phi_2. Phi one with psi two the side you subtract this portion out and you normalize the whole thing. Now that ensures this vector which I now call phi_2 is normalized to unity and moreover is orthogonal to phi_1.

(Refer Slide Time: 00:27:51 min)



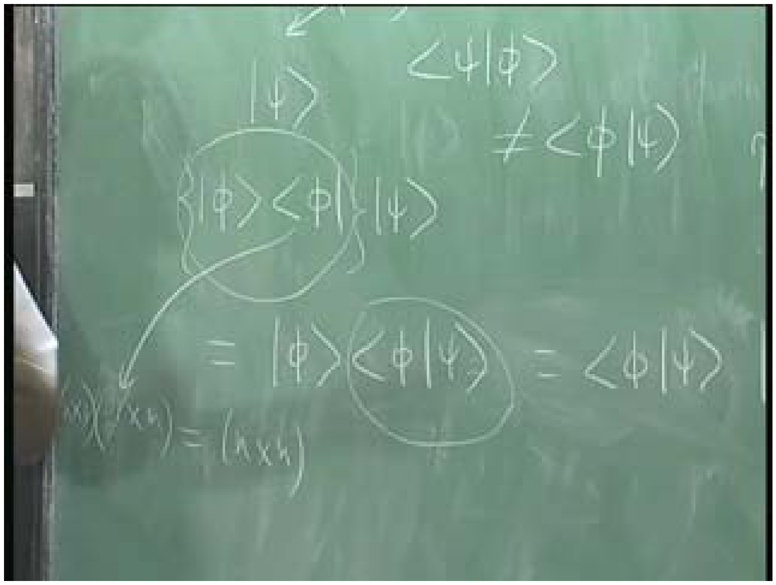
So this will guarantee that phi$_1$ phi$_1$= 1= phi$_2$ phi$_2$ and moreover phi$_1$ phi $_2$ = 0. Now once I got phi$_1$ and phi$_2$, I take psi$_3$ and subtract from it the component of psi$_3$ along both phi$_1$ and phi$_2$ and I normalize the whole thing. So in this systematic way I end up with a set of vectors phi$_1$, phi$_2$, etc which would really be the original vectors psi$_1$, psi$_2$ and so on with portion subtracted such that it forms an orthonormal basis. Notice what has gone on here (Refer Slide Time: 28:53) in this case. What I did was to say psi $_2$ is a vector. Here was psi$_1$ symbolically and this has some arbitrary length. So I divide by its norm and then I created this vector phi $_1$.

(Refer Slide Time: 00:29:08 min)



This is phi$_1$ by dividing by its length that's a unite vector now. And then I had a psi$_2$ and I took the dot product of psi$_2$ with phi$_1$ on the left hand side. That essentially is this portion and I subtracted that out. So what was left out was this vector (Refer Slide Time: 29:45) and then I normalize that vector to unity. So this became phi$_2$ and so on. So this is all I did in this orthogonalization procedure. What you have to note is that to remove the portion of this vector (Refer Slide Time: 30:01) along the vector phi$_1$, what I did was to take a dot product with phi$_1$ on the left hand side and then I multiply by the unit vector phi$_1$here.

(Refer Slide Time: 00:30:19 min)



So what I have tacitly done is to use the fact that if you give me any vector psi and I take any other unit vector phi then this object here applied to this vector psi (Refer Slide Time: 30:40) from the left projects that portion of psi which is along the unit vector phi. So if I act on the vector psi from the right hand side, this is nothing but phi by psi which is just a complex number. You can remove it to the left or right. So it's a coefficient multiplied by unit vector here.

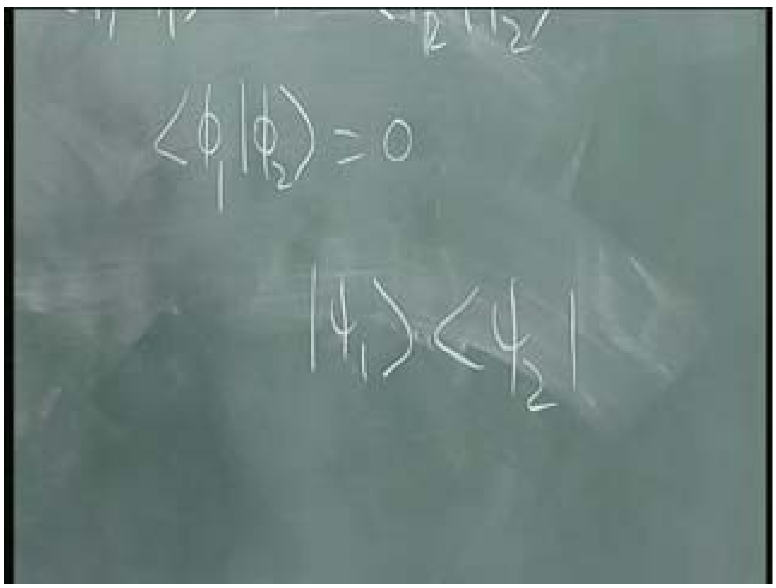
Given a set psi_1, psi_2, etc I would like to create an orthonormal set and I called that set phi_1, phi_2, etc. What I am given is a set of vectors which are not perpendicular to each other. What I am creating is a set which is perpendicular to each other. I will continue to use phi as far as possible for orthonormal base sets. And the reason is I am going to use symbol phi for arbitrary state vectors of quantum mechanical systems. So I don't want to confuse it with basis vectors. The point I made here was that these objects are different objects. This vector should be identified with a column vector of some kind. This is not a column vector (Refer Slide Time: 33:40).

This object is a complex number and whether you write column vector times number or number times column vector doesn't matter and therefore I moved it here. On the other hand, if I identify this with the a column vector and this with a row vector this multiplication here is column on the left row on the right. So what sort of object is this? It's a matrix because if this is represented by a column vector with n rows and one column then this psi here (Refer Slide Time: 34:43) is an n by one matrix. Its n rows and one column. This quantity here (Refer Slide Time: 34:58) on the other hand is a one by n and this here is n times one. So this is an n by n object this goes away and it's an n by n matrix and it acts on a column vector to produce a column vector once again in the left.

So objects of this kind (Refer Slide Time: 35:20) are to be identified with operators. They act on vectors and produce other vectors. So this gives us a first introduction to the very important idea that in a linear vector space you have, in addition to the vectors, another set of objects called operators and these operators would act on the vectors and produce other vectors.

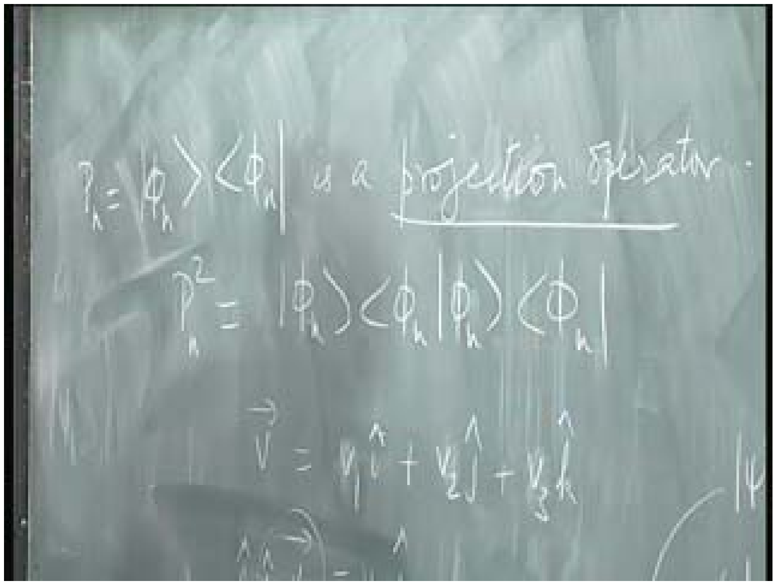
A good way of remembering how beautiful this notation is ket vectors are like column vectors, bra vectors are like row vectors and if you put bra on the left and the ket on the right, you get a number but if you put the ket on the left in the bra on the right you get an operator. Now of course it does in a given linear vector space even this (Refer Slide Time: 36:15) is an operator.

(Refer Slide Time: 00:36:17 min)



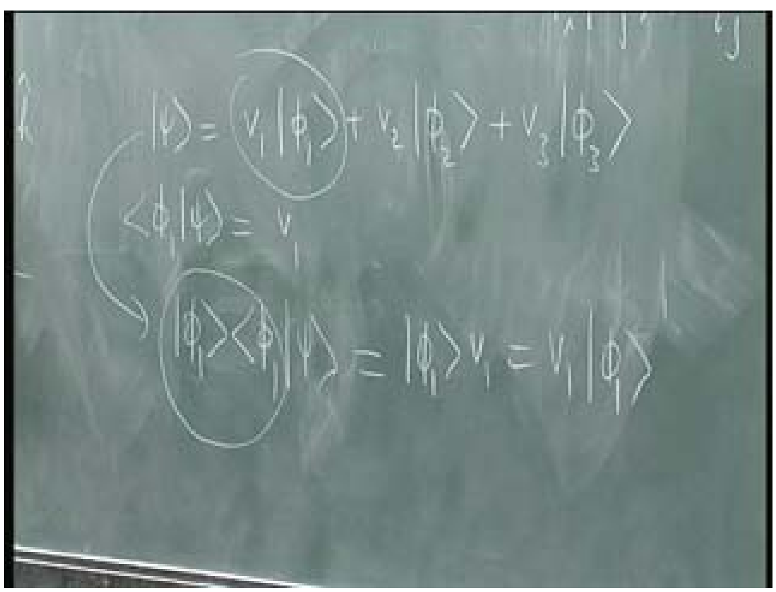
Psi $_1$ psi $_2$ does not have to be the same on both sides in a given vector space with a certain dimensionality. You have to take a ket vector and take a bra vector and if the spaces are same dimensions, they make an n by n object like a matrix. so such quantities are operators but in the special case in which psi$_1$= psi$_2$= phi, this operator is a very special operator because you have a ket vector and the same bra vector on this side and when it acts on other ket vectors, it projects out the portion of that ket vector along this vector. So it’s called a projection operator.

(Refer Slide Time: 00:37:20 min)



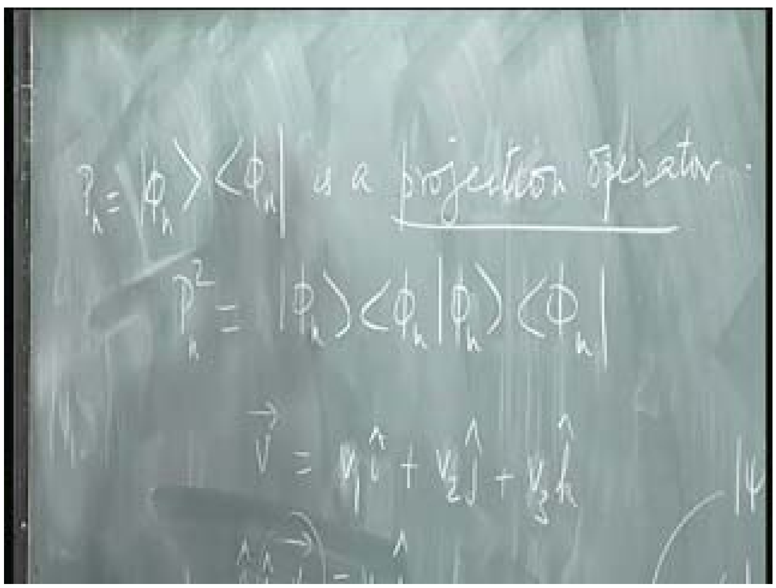
So this object phi $_n$ phi $_n$ is a projection. That's what you do when you take an arbitrary vector in three dimensional spaces. Let me write it in terms of the familiar notation i,j,k. If I have $V = V_1$ i + V_2 j + V_3 k, how do I find the coefficients V_1, V_2 and V_3? I take i dot the vector because I know this is an orthonormal basis. So I dot j= 0 and I dot k =0 and what I am really doing is to take i dot V on the left hand side and that's guaranteed to give me V_1.

(Refer Slide Time: 00:38:36 min)



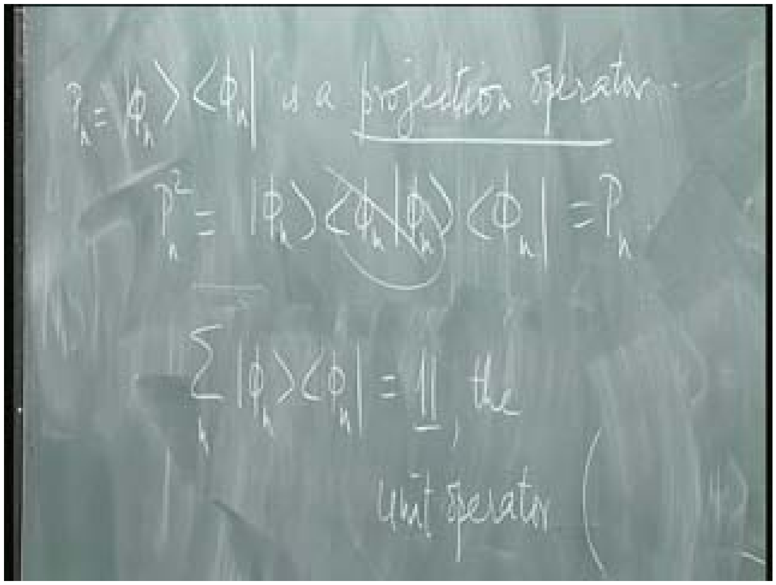
In abstract notation the same thing would look like some vectors psi = V_1 phi $_1$+ V_2 phi $_2$ + V_3 phi $_3$ and I want to find out what's V_1 here and given the orthonormality relation phi$_i$ phi $_j$= delta ij and this is an orthonormal basis. To find V_1, I have to take the dot scalar product with phi$_1$ and psi which would give V_1 because phi$_1$ and phi$_2$ is 0 and phi$_1$ and phi$_3$ is 0. Normally we are used to calling components of vectors as vectors themselves which is not true. V_1, V_2 and V_3 cannot be called as components of the vectors since components mean a part of the vectors and the parts of a vector can only be vectors. So I should I call the entire V_1 phi $_1$ and so on as components. To get the components, I do phi$_1$ phi$_1$acting on the vector psi. And what does this give you? Well when phi $_1$ hits this, it gives you one. When phi$_1$ hits psi, it gives a V_1 and what we are left with is a ket phi$_1$. So we really have phi $_1$ V_1 which is the same as V_1 phi $_1$. So the projection operator is this (Refer Slide Time: 40:45). In three dimensional vector space, very often in addition to dot product and a cross product, in the olden days there used to be quantities called dyadic or tensor products.

(Refer Slide Time: 00:40:58 min)



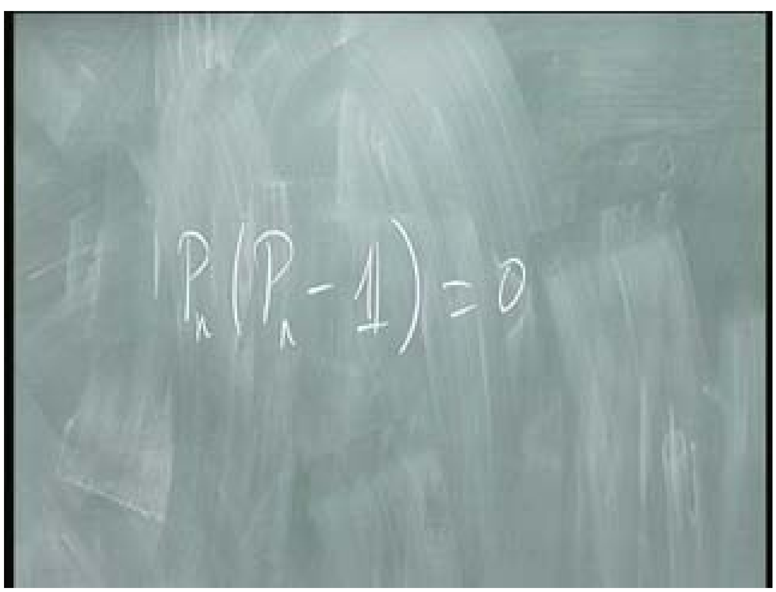
So they would actually write the projection operator as i i. the idea is that if you took ii and doted with V on the right hand side, you get V_1 i and of course you immediately see that once I take up a vector and project along all directions, I get the vector itself. So it is obvious that this is a projection operator and it has following properties. Let's call the projection operator as P_n. so P_n squared will be twice the operator. Once you take a vector and project along the x axis and you project again, you are just going to get nothing new. It's going to be itself. P_n squared should be P_n itself and indeed that is true because all you have to do is write it twice together.

(Refer Slide Time: 00:42:47 min)



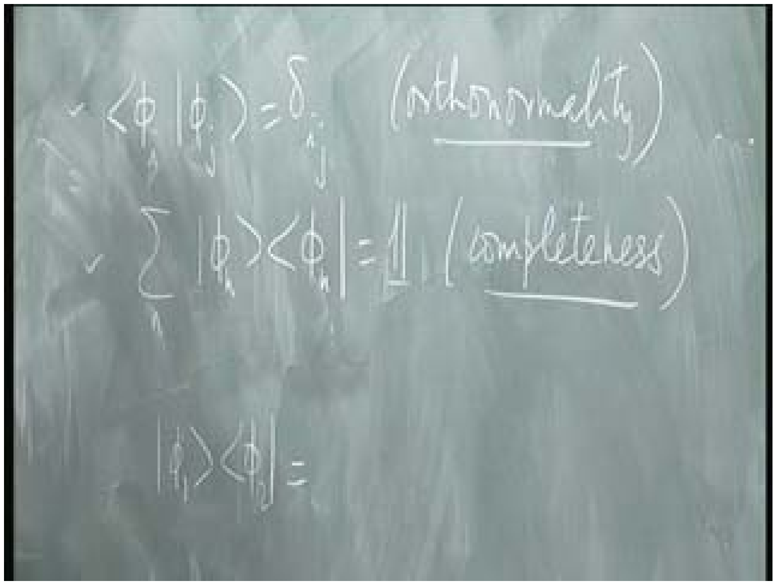
What happens if you took a summation over phi $_n$ phi $_n$? This is going to produce some of all the components of this vector. In other words it's going to produce the vector itself. So this is equal to the unit operator. The unit operator is something which when applied to a vector reproduces the vector.

(Refer Slide Time: 00:44:07 min)



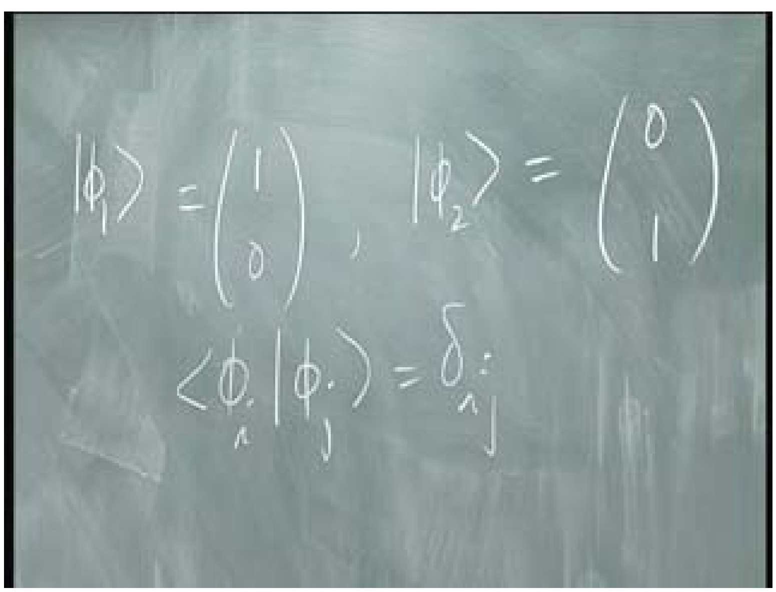
Another interesting property is for example, Pn squared is Pn. So Pn (Pn – 1) = 0. Imagine Pn is an some finite dimensional vector space and matrix squared is equal to the matrix itself, such a matrix is called as the idempotent matrix where if you square it, you get itself. So the Eigen values of the projection operator are 0 or 1. So we are going to use two very crucial properties of orthonormal basis in linear vector spaces

(Refer Slide Time: 00:45:43 min)



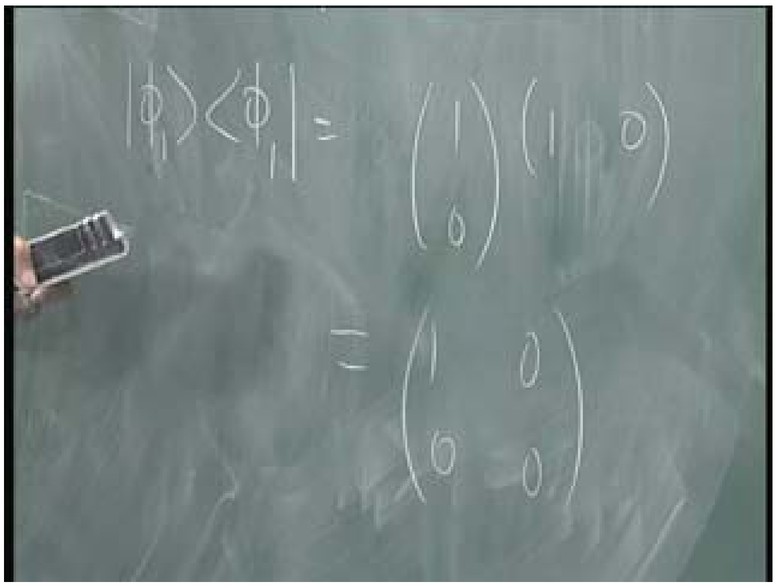
The first one is orthonormality and the second one is a set of projection operators and this is called completeness. For example let's look at a two dimension linear vector space.

(Refer Slide Time: 00:47:07 min)



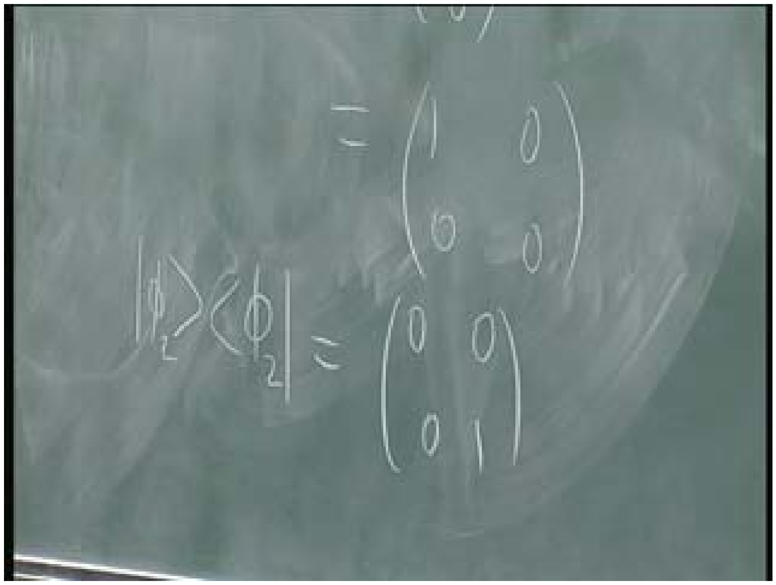
In this two dimension vector space, I would like to represent the unit vectors along the x and y directions by column vectors. This is a representation on the right hand side this is the abstract vector. In a two dimensional space, the natural thing to do is to use column vectors with two rows and one column. phi$_2$ is zero one and that's a y direction. It's immediately clear that the inner product of phi$_1$ with phi$_2$ is 0 and phi $_1$ phi $_1$ is 1. So phi $_i$ phi $_j$ = delta $_{ij}$ is satisfied. Now let's ask what the projectors look like.

(Refer Slide Time: 00:48:08 min)



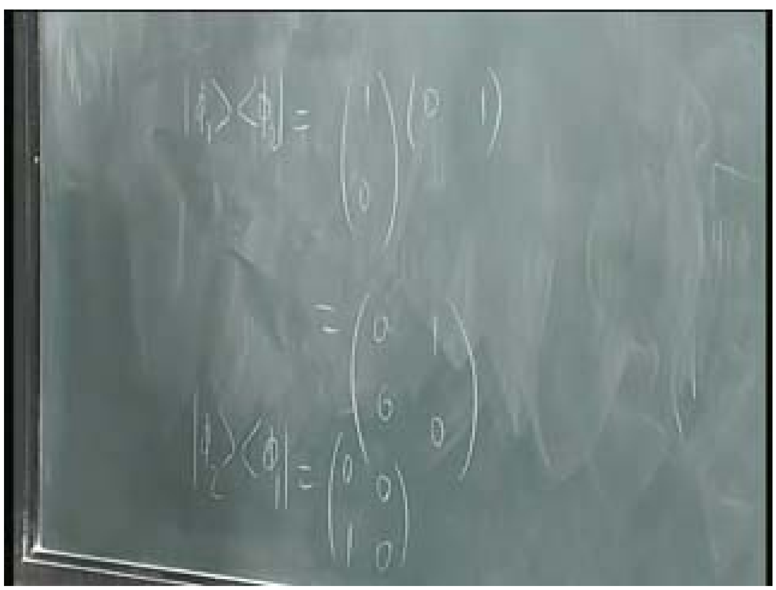
Now phi 1with phi 1 has to be a matrix. This would give a 1, 0 0, 0. You can also have phi $_2$ with phi $_2$ which would result in the matrix 0, 0 0, 1.

(Refer Slide Time: 00:49:07 min)



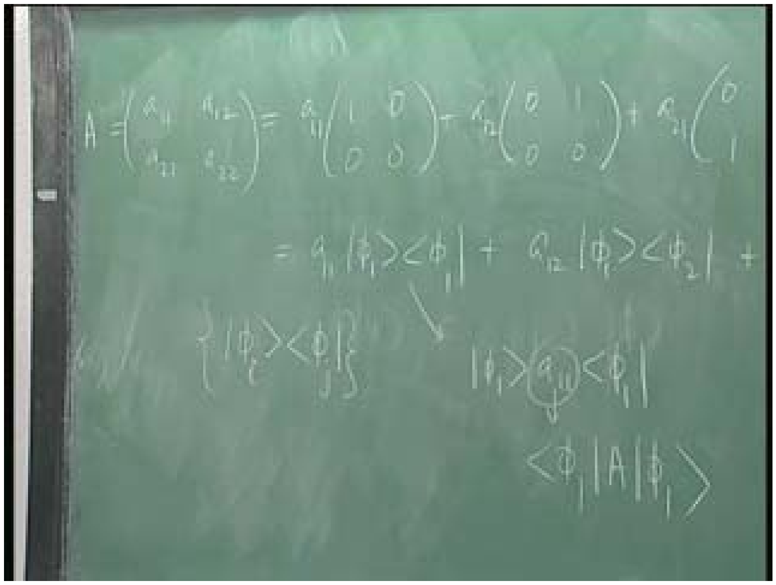
So this satisfies the completeness rule since the sum of phi_1 phi_1 and phi_2 phi_2 gives you a unit matrix.

(Refer Slide Time: 00:49:31 min)



Phi $_1$ with phi $_2$ will be a 2 by 2 matrix. This is also an operator but not a projection operator. This is equal to 0, 1 0, 0. Now phi_2 with phi_1 will give a 0, 0 1, 0.

(Refer Slide Time: 00:50:38 min)

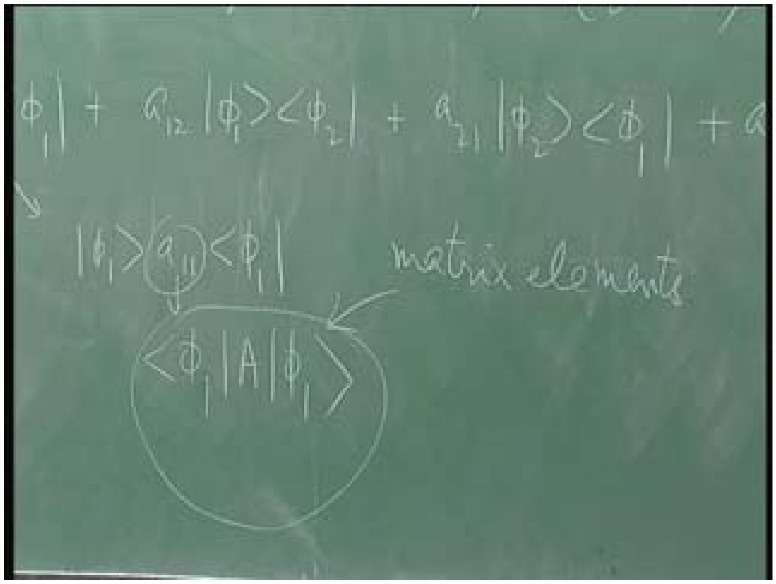


Now you know that any 2 by 2 matrix in the natural basis, if I write a b c d, this represents an operator because it acts on column vectors and produces another column vector. So it's a general operator but this general operator is shorthand. This is shorthand for a times 1,0,0,0 + b times 0,1,0,0 + c times 0, 0, 1, 0 + d times 0,0,0,1. So it's clear that these four matrices are forming a basis in the space of operators which act on the vectors of your original two dimensional linear vector space. Let us denote a as a_{11}, b as a_{12},c as a_{21} and d as a_{22}. So it's a_{11} phi_1 phi_1 + a_{12} phi_1 phi_2 + a $_{21}$ phi_2 phi_1+ $a_{22}$$phi_2$ phi_2.

This is an explicit matrix representation and these are abstract operators (Refer Slide Time: 52:45). When I provide a basis of this kind an orthonormal basis not only can I write every vector as a linear combination of these unit vectors but I can also represent operators on these vectors in the natural basis which is the set phi $_i$ phi $_j$. Just as the set of vectors phi i or phi j, the ket vectors form a basis. The set of operators ket phi $_i$ bra phi $_j$ where i and j run over all the possible values. They form a basis for expanding operators. Now you begin to see why a_{11} is called the matrix element. The reason is you could write this term as phi_1 a_{11} $phi_{1.}$ How do I find a_{11} given an arbitrary vector psi? The way I find its component along any of the unit directions is to use the projection operator. So $a_{11}$is nothing but bra phi $_1$ A ket phi $_1$. So if I want to find a_{11} from this relation, I do ket phi 1 on the right. That kills this (Refer Slide Time: 55:49). I do bra phi_1 on the left. That kills this because phi 1with phi1 is 1. And you get a $_{11.}$ Then phi_2 with phi_1 is 0. So that goes away. phi $_1$ with phi $_2$ is zero and here both are zero. So now we begin to see that this notation is so beautiful that ket vectors represent vectors in the space and bra vectors represent vectors in the dual space. a bra on the left and ket on the right is an inner product. It's a complex number. a ket on the left, bra on the right is an operator. The basis in the linear vector space also provides basis for operators in this space.

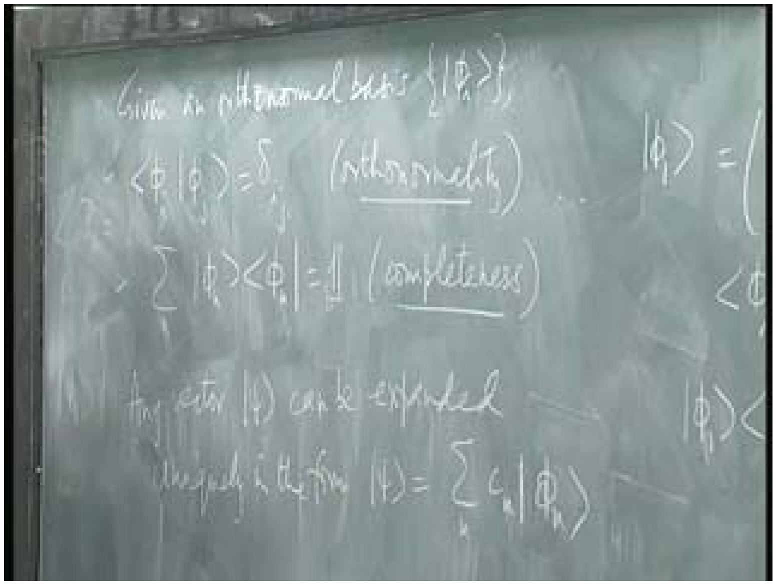
And finally objects like this (Refer Slide Time: 56:18) are the matrix elements. So these quantities are called matrix elements.

(Refer Slide Time: 00:56:45 min)



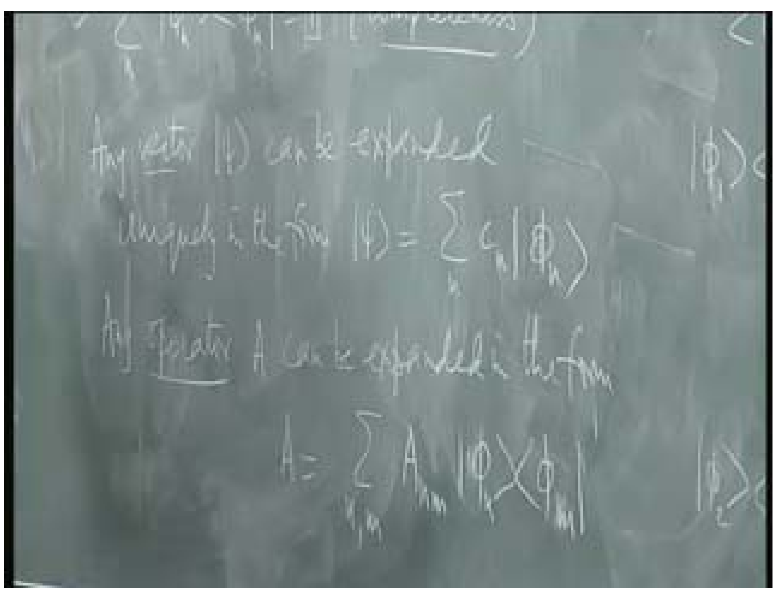
And I will continue to use the work matrix element even when the vector space is infinite dimensional and these operators are not matrix operator. They may be differential operators, integral operators, integro differential operators and anything at all. We are going to look at spaces where there number of dimensions is not only infinite dimensional but continuously infinity dimensional. So you can’t even label letter one two three four up to infinity but this is goes to continuously. So the concepts are not very hard to generalize. We have to be careful of some technicalities like we were about $l_2$and so on but I would continue to use these objects as matrix elements. Now when we come to the postulates of quantum mechanics then we will see that objects like this are measurable quantities. These are the measured values. These are the things that you would get to make actual physical measurements. So it’s important to recognize what sort of objects we have here.

(Refer Slide Time: 00:58:11 min)



So the third lesson is writing here given an orthonormal basis phi $_i$, you first have this then you have completeness and then you also find that any vector psi can be expanded uniquely in a form psi = summation over n C_n phi$_n$ uniquely. That's the beauty of these orthonormal basis. that once you make an expansion then the set of numbers C_n uniquely specifies the vectors psi and vice versa.

(Refer Slide Time: 00:59:16 min)



Any operator A in this space just as any vector can be expanded in the form A = summation over n, m A $_{nm}$ phi $_{n}$ phi $_{m}$ and these are called the matrix elements of this abstract operator. So this (Refer Slide Time: 01:00:13) set forms a basis. Of course when n= m, these are diagonal matrix elements and when n is not equal to m, these are off diagonal elements. There is no guarantee here that every matrix can be written as diagonal form. We will talk about Eigen values of operators and so on very shortly but this is the basic mathematical framework that you need. We need a few more important concepts such as what is meant by Hermitian conjugate, an adjoint and so on. We will talk about that next time but I want to impress upon you very firmly the fact that this machinery, once you read it now it's almost automatic which is a most self correcting. So completeness and orthonormality are very useful relations. The next thing is to talk about function spaces and I will do that tomorrow.