FOUNDATIONS OF QUANTUM THEORY: NON-RELATIVISTIC APPROACH

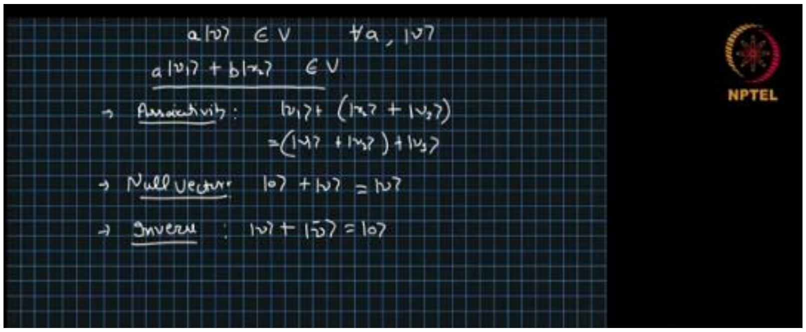
Dr. Sandeep K. Goyal
Department of Physical Sciences
IISER Mohali
Week-01
Lecture-03
Linear Vector Space

Hello everyone. In this lecture and next few lectures, we will be discussing the mathematical prerequisites required for the foundations of quantum mechanics course. So, in this series on the set of lectures we will cover the linear vector spaces, linear operators and the representation, matrix representations and stuff like that. So, throughout this course, we will come back to mathematical methods, mathematical prerequisites in which we will cover the tensor operators and whatever is required mathematically to understand this course. So today we will start with linear vector spaces.

A linear vector space V is a set of vectors we represent by $\{|v_i\rangle\}$, which satisfies certain properties. Those properties are the following there is a vector space set of vectors over a field of scalars, the set of scalars let us call them $\{a_i\}$, these are vectors these are scalars so a vector, a set of vectors $\{|v_i\rangle\}$ over a field $\{a_i\}$, is called linear vector space if it is closed under vector addition and scalar multiplication. What we mean by that is if we have a vector $|v\rangle$ and we multiply it with a scalar a, then also it's a vector from the same vector space in that. The set v should contain, if it contains $|v\rangle$ then it should all contain a$|v\rangle$ for every a and for every $|v\rangle$ okay, not just that if we have two vectors $|v_1\rangle$ and $|v_2\rangle$, then there sum, which sum is some abstract operation we have defined over this vector space so the sum should also be from the same vector space, same set. So we can write one statement which contains both the conditions $a|v_1\rangle + \mathrm{b}|v_2\rangle$ should belong to the same set so this is what we mean by closed under vector addition and scalar multiplication.

Further, associativity, if we have three vectors $|v_1\rangle, |v_2\rangle$ and $|v_3\rangle$, then it shouldn't matter in what order we add them we should get the same outcome. Then, there is a requirement of null vector. There should exist a vector $|0\rangle$, we call it zero, such that if we add it with any vector $|v\rangle$, the resultant is the vector $|v\rangle$. So $|0\rangle$ should also exist in this set. And if there is a null, there will be some inverse. The inverse vector for a given vector $|v\rangle$, the inverse will be defined by $|\bar{v}\rangle$ such that $|v\rangle + |\bar{v}\rangle$ is the null vector. So, if we have a set of vectors, vectors can be any abstract quantity or any abstract mathematical structure such that it can be worked out with the scalars and we can define operations such as scalar multiplication and vector addition over that then that that set of vectors can be called a linear vector space if it satisfies these four conditions, okay.

(Refer slide time: 3:51)



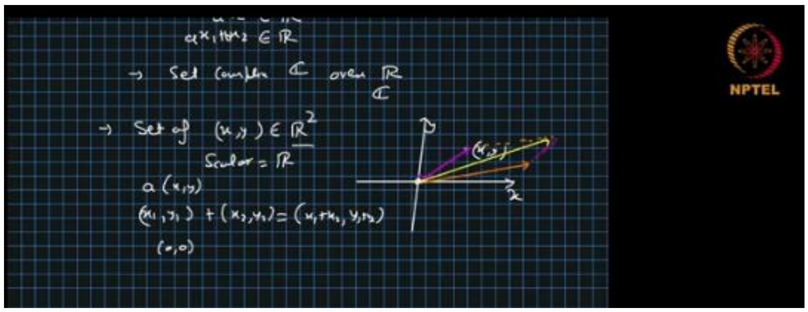
So, there is a little bit more condition, but these are the main four conditions which need to be satisfied. And if that satisfies, then we can call it a linear vector space. It has certain property which will be revealed as we go along this lecture. And we will know eventually this can be very powerful mathematical structure to understand many concepts in foundations of quantum mechanics. So to understand it better, let's take examples.

So, set of all the real numbers be represented by this R. And this is a scaler, this set is, it forms a linear vector space. We have to define this; it forms a linear vector space over the field of real numbers. Okay. So, if we have a x, which is a real number from this set then we multiply it with the scalar which is also a real number that also belongs to, ax also belong to R. So, if we have x_1 plus x_2 or ax_1+bx_2, this is also a real number, so it belongs to R. So, it satisfies the first condition that is closed under vector addition and scalar multiplication Now, if we take three real numbers, it does not matter how we add them.

If we take the definition of addition as the regular definition of addition, then it does not matter in what order we add them, we get the same number. So, it satisfies the associativity. There is a null vector that is the real number 0, such that if we add 0 with any number, we get the same number back. So that's the null vector. And for every real number x, there exists a -x.

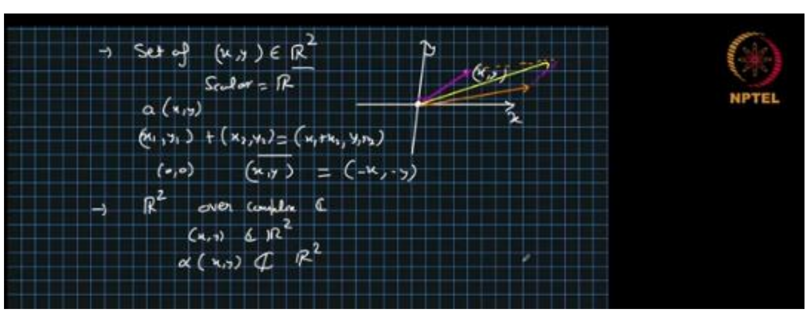
So, such that x+(-x) gives us the 0, so, we get a inverse also. So, set of real numbers also form a vector space over real field. Similarly, we can have set of complex numbers we represented by C, over real field. It forms a vector space over complex field, it forms a vector space. It's trivial to understand all this thing, we can do it ourselves. Slightly more non-trivial set example will be set of, ordered set of two real numbers x and y. So, this is represented by R^2, so this we know, it is the two-dimensional plane we can have. So, any vector in this two-dimensional, any point in this two-dimensional vector, two-dimensional plane is represented by x, y. So, that is the geometrical representation of R^2. Now, the scalars will be the real numbers. If we multiply vector x, y with a scalar, we get another vector in the same two-dimensional plane.

(Refer slide time: 8:45)



If we have, we can define the addition operation as x_1+x_2, y_1+y_2, so that will would amount to having one vector, second vector, so, and the sum will be given by this vector. So, that is the vector addition we have defined in R^2. So, under this vector addition we can see that it's closed over scalar multiplication and vector addition because addition of two such vectors will give us another vector in the same R^2. Then there is associativity, the three vectors we add in any form any way it does not matter we will get the same answer. And the null vector will be the (0,0), that is the origin of this two-dimensional plane. And invert will be, (x, y) inverse will be (-x, -y). So, in that way, it satisfies, R^2 satisfy all the properties required to be linear vector space. It's not enough to give the examples in favor.

(Refer slide time: 9:41)

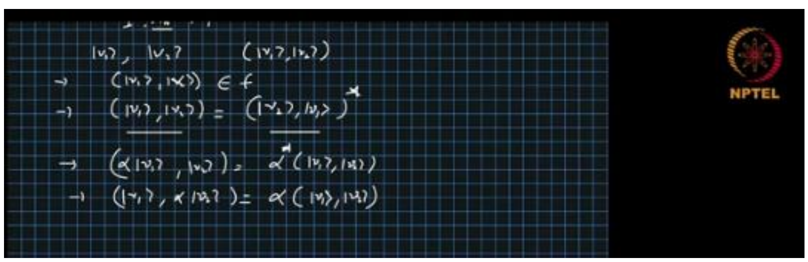


To understand something, it's also important to give some examples which do not satisfy the criteria, for example, we can have R^2 over complex field C, then it does not form a vector space. Because if we have (x, y) which is from R^2, then, if we multiply with α, some complex number, then it does not belong to R^2 because it's no longer real. So, in that way, this is one simple example where we see that a set of vectors does not satisfy the linear vector space criteria. So R^2 over complex is not a linear vector space. Another example which can be very interesting, which will be very useful for us is set of n by m matrices, set of all the n by m matrices, complex matrices over complex field is also vector space. If we take a matrix M, which is from this set, then M_1+M_2 will also be the same n by m matrix complex matrix so if we multiply it with α_1 and α_2, then also it's from the same vectors, same set. So, it's all, if we check one by one all the conditions, we will see that uh for example the null matrix will be the matrix of zeros is the null matrix and minus of M will be the inverse matrix or inverse. So, in that way, we will see that this set of n by m complex matrices also form a linear vector space. One of the very interesting subsets of this is of particular interest to us, that is the Hermitian operators, Hermitian matrices. Hermitian matrices H let us say is defined as a matrix which is equal to its own Hermitian conjugate. Hermitian conjugate is defined as transpose and conjugate. So, if a matrix is equal to its own Hermitian conjugate, then it's the Hermitian matrix, and this is interesting from quantum mechanics point of view because all the observables are represented by Hermitian operators, Hermitian matrices.

So, this matrix, this set of matrices, Hermitian matrices form a linear vector space over real field. Although the elements of the Hermitian matrix can be complex numbers but it forms a linear vector space only over the real field not on the complex field. So, if we have matrix H_1 which is Hermitian and we have H_2, which is also Hermitian then, and we multiply with scalar field a and b, if we take the Hermitian conjugate of that it will be $aH_1^\dagger + bH_2^\dagger$, which is $aH_1 + bH_2$, so it's the same matrix so it's a Hermitian. So, in that way it's closed over scalar multiplication and vector addition. Associativity, this is the simple addition of the matrices so it does not matter in what order we add them so associativity is trivially satisfied. The null vector, the matrix of all the zero elements is a Hermitian matrix so that will be the null vector and a negative of a Hermitian matrix is also a Hermitian matrix so that will serve as the inverse vector. So, in that way the set of all the Hermitian matrices over real field forms a linear vector space. And this will be used extensively in the entire course. So, with these examples, I hope we have a little bit of better understanding of what it takes to be a linear vector space but we do not stop here at the definition of linear vector space, we can introduce further structure over the vector space that is for example we can introduce inner product.

Till now we had any operation involving two vectors was just the addition which was the vector addition but now we can introduce another structure that is the inner product between the two vectors, so if you have an inner product, is a mapping on the vector space, the inner product, let me represent by I, which takes vectors of V and it's mapped to scalar. Scalar field is generally represented by f, so v goes to f okay so it's actually v x v, so two vectors go to f. So, we can define, if we have two vectors $|v_1\rangle$ and $|v_2\rangle$ from the vector space V, then we write the inner product as $(|v_1\rangle,|v_2\rangle)$, okay so ordered set of these two vectors and it satisfies certain property. First of all, the property is that $(|v_1\rangle,|v_2\rangle)$, which is a scalar, $(|v_1\rangle,|v_2\rangle)$, is equal to $(|v_2\rangle,|v_1\rangle)^*$, it's a conjugate. The conjugate in any sense we define conjugate on the field. So, this will be one element of the field and it will be the conjugate element of the field. Typically, the kind of vectors we are interested in, the field is complex. So, then it will be a complex conjugate.

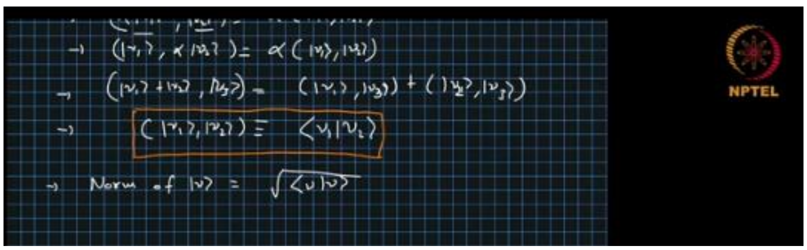
(Refer slide time : 16:55)



So, the inner product of two vectors will give you some complex number and the inner product of the same two vector, but in the reverse order will give us the complex conjugate of the same number. Now, if we have a scalar times the first vector $|v_1\rangle$ and $|v_2\rangle$, the second vector, then it will be $\alpha^*(|v_1\rangle,|v_2\rangle)$, and if we have $(|v_1\rangle, \alpha|v_2\rangle)$, then it will be $\alpha(|v_1\rangle,|v_2\rangle)$. It means the inner product is linear in the second argument and antilinear in the first argument okay and then there are some simple relations $|v_1\rangle,|v_2\rangle,|v_3\rangle$ with three vectors such that we are taking the inner product of $|v_3\rangle$ with the sum of $|v_1\rangle$ and $|v_2\rangle$, then, it can be decomposed as $(|v_1\rangle,|v_3\rangle)+(|v_2\rangle,|v_3\rangle)$.

It seems very trivial but sometimes these properties can be very important. So, throughout the quantum mechanics course or foundations of quantum mechanics course, we will be representing $(|v_1\rangle, |v_2\rangle)$, as $\langle v_1|v_2\rangle$, so this will be the definition, this is how we will be representing the inner product. Once the inner product is defined over a vector space, then certain structures or certain definitions can be given, for example, we can define the norm of a vector, that is given by $|v\rangle$ with the inner product of it with itself, and the square root of that, that is this is how we define the norm which is basically the length of the vector in some metric and we can sure like from the first definition of the inner product we know that $\langle v|v\rangle$ is a real number, not just that $\langle v|v\rangle$ is also a positive number okay and equality holds only for the null vector, not generally every vector will have non-zero no other than the null vector. We can define the orthogonal vectors, that is you have two vectors, $|v_1\rangle\ and\ |v_2\rangle$, they are orthogonal if their inner product is zero then they are called orthogonal, normalized vectors, that if the norm of a vector is 1, then it is called a normalized vector.

(Refer slide time: 17:19)

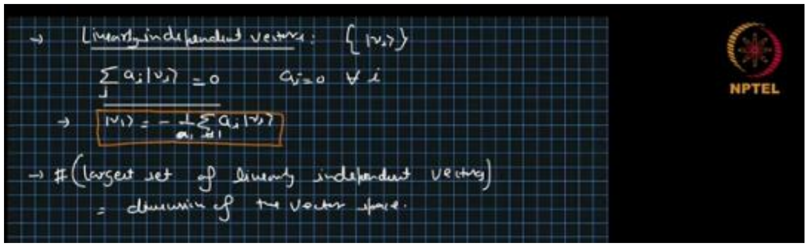


So, if we have linearly independent vectors, if we have a set of vectors $|v_i\rangle$, then they are called linearly independent, if $\sum_i a_i|v_i\rangle = 0$, this equation has the only solution such that a_i is 0 for all i, then only, we call this set of vectors linearly independent. What does that mean? Let us say there exists a solution where not all a_i are zero and we get this, we can satisfy this equation by taking a set of a_i such that not all of them are zero. So, in that case, we can represent, let us say $|v_1\rangle$ as $-(1/a_1)\sum_{i\neq 1} a_i|v_i\rangle$. So, it means one vector out of these all vectors can be written as the sum of all the other vectors.

So, in that way there is a, so let us say a_1 was non-zero. So, it means we can write the vector $|v_1\rangle$ as a linear sum of ll the other vectors in this set, some of these other vectors in the set. In that way, one vector depends on the other vectors. So, in that sense, it is not independent set.

So, the independent set will be the one in which none of the vector can be written as a sum of the other vectors. Okay, so, what is the significance of this, if we get the largest set of linearly independent vectors in a vector space, okay, then the cardinality of this set, cardinality of this called the dimension of the vector space. The cardinality or the number of, the maximum number of linearly independent vectors in a vector space tells us the dimension of the vector space. What is the dimension of the vector space? For R^1, the dimension was 1 because you have just one number which is, which is one vector which is independent and all other vector depends on it if you just multiply it with scalars.

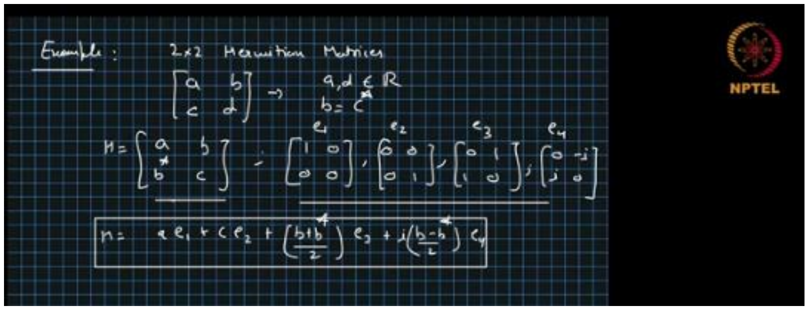
(Refer slide time: 21:31)



In R^2, we need at least two vectors to represent all the other vectors. So, in R^3, it would be three in different vector spaces, we have different number of vectors. So, for example, and this I will insist that you verify if we have two by two Hermitian operators. We said that Hermitian matrices form a vector space. So, we consider 2 by 2 Hermitian operators.

Then we can write it as some $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$. And we can see if we put the Hermiticity condition, then it will give us a and d in real numbers and $b = c^*$. So, we can write an arbitrary Hermitian operator as $\begin{bmatrix} a & b \\ b^* & d \end{bmatrix}$. Now how many, what is the dimension of this vector space, we need to ask and I'm giving you the example or I'm giving you the answer but please verify this thing. We can choose vectors, this is one vector we can choose another vector, this is a Hermitian operator. We can choose another vector which is Hermitian operator. We can choose another $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and we can choose $\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$. These four matrices are the Hermitian operators are some Hermitian operators and you can write any of the operators of this form as a linear sum of these four operators with real coefficients. For example, we can write H to be a times, let me call it e1, e2, e3, e4, the vectors. Then I can write H as $ae_1 + ce_2 + ((b + b^*)/2)e_3 + i((b - b^*)/2)e_4$. So, in that way, we can write an arbitrary matrix as a linear sum of these numbers. So we can write an arbitrary matrix H Hermitian operator as a linear sum of these four operators. So it means if I have a set of five Hermitian operators such that four of them are e_1, e_2, e_3, e_4, and the fifth one is something else. Then I can write that matrix, fifth matrix is the sum of the four, e_1, e_2, e_3, e_4.

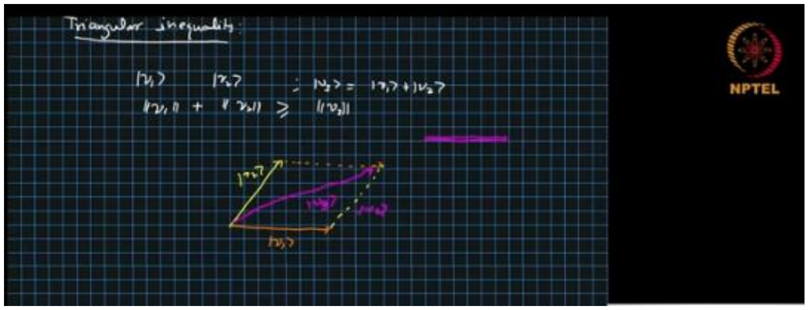
(Refer slide time: 25:02)



So, the largest independent, the largest set of linearly independent vectors in the linear vector space of Hermitian operators will have the cardinality of four because you can have only four independent vectors in it. So, it means the dimension of two-by-two Hermitian operators is four, okay, so dimension of such vector space is four. So, later on we will be using this definition, the definition of these dimensions to characterize many vector spaces and to talk about few things which will be clear soon. So, once we define the inner product structure or the linear vector space, there are many interesting properties which emerges out of these structures. One of them, one of the most useful is the triangular inequality. So, it goes like this. If we have two vectors, $|v_1\rangle$ and $|v_2\rangle$, then their norm $||v_1||$ and $||v_2||$ that is their lengths in some sense, the sum of their length is always greater than or equal to the sum of the length v_3, where $|v_3\rangle$ vector is the sum of $|v_1\rangle$+$|v_2\rangle$. So, the sum of the individual length of two vectors is more than the length of the sum of the vectors. So, we can graphically, we can see it like this, we have a vector $|v_1\rangle$ and a vector $|v_2\rangle$, then there's sum, as we know from our earlier experience the sum is represented by $|v_3\rangle$. So, in this, the three vectors $|v_1\rangle$, $|v_2\rangle$, this is $|v_2\rangle$ and $|v_3\rangle$, they form a triangle. So, basically what we are saying in this inequality is the sum of two sides of a triangle is always greater than the third side.

It can be only equal when it is not really a triangle, when we have two vectors $|v_1\rangle$ and $|v_2\rangle$ collinear and third one is also in the same direction. So that's the trivial case, but if we have a valid triangle, then the sum of the two sides is always greater than the third side. So, this inequality is true in general for any vector space where we have defined the inner product using the definitions we discussed earlier. So, to prove it, we start with something called Cauchy Schwarz inequality. This is another very interesting inequality in linear vector space with inner products.

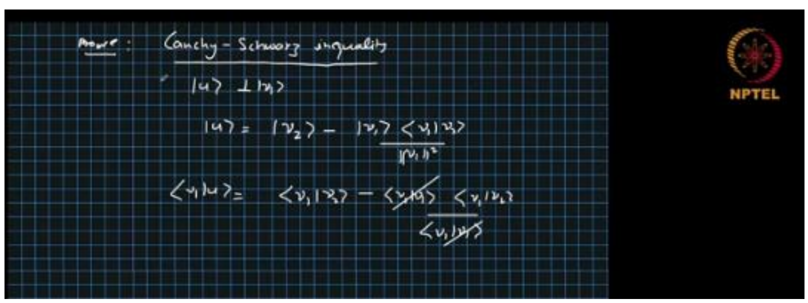
(Refer slide time: 27:54)



So, this goes like this. If we, let us define a vector $|u\rangle$, which is orthogonal to $|v_1\rangle$. So, how do we get this thing? This is orthogonal to $|v_1\rangle$ and it is in the same plane as $|v_1\rangle$ and $|v_2\rangle$. So, how we define it, then we can say $|u\rangle$ is

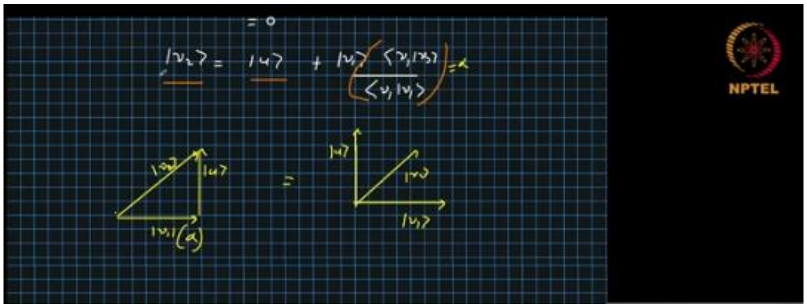
$|v_2\rangle$-($|v_1\rangle\, \langle v_1|v_2\rangle)/||v_1||^2$. Now, if we see the inner product of $|v_1\rangle$ with$|u\rangle$, that will be $\langle v_1|v_2\rangle - \langle v_1|v_1\rangle\langle v_1|v_2\rangle/\langle v_1|v_1\rangle$. This cancels and we have $\langle v_1|v_1\rangle$ and we have $\langle v_1|v_1\rangle$ and that's zero so it means $|v_1\rangle$ and $|u\rangle$ are orthogonal.

(Refer slide time: 29:43)



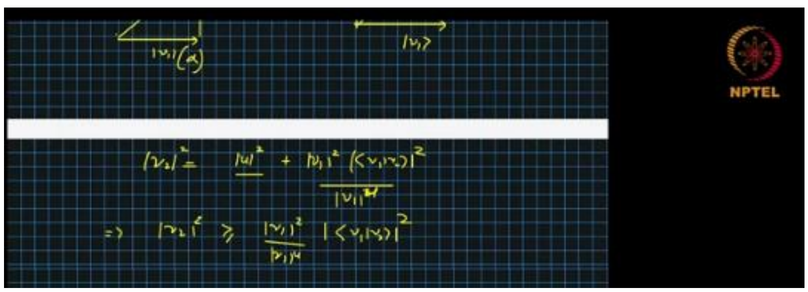
Now we can write $|v_2\rangle$ as $|u\rangle + |v_1\rangle\langle v_1|v_2\rangle/\langle v_1|v_1\rangle$. So, from here we can see that $\langle v_1|v_2\rangle$ is a sum of two orthogonal vector $|u\rangle$ and $|v_1\rangle$ and rest is just a scalar, this thing is just a scalar. So, $|v_2\rangle$ can be written as two orthogonal vector and if we want to write it graphically, if we want to represent it graphically, we have $|v_1\rangle$, we have $|v_2\rangle$ and we have $|u\rangle$, and we are writing $|v_2\rangle$ in terms of $|u\rangle$ and $|v_1\rangle$ with appropriate weights. So, if we represent it in a slightly different way, it will be $|v_1\rangle$ with some scalar times $|v_1\rangle$, α, where this is α and this is $|u\rangle$ and this will be our $|v_2\rangle$.

(Refer slide time: 31:16)

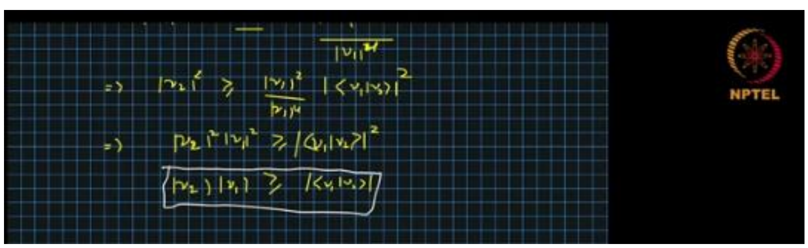


So, we have $|v_1\rangle$, $|v_2\rangle$ and $|u\rangle$ forming a right-angle triangle. So, now we can use in a way the Pythagoras theorem or we can just take the, we can, from here we can see that the norm of $|v_2\rangle$ will be the norm of $|u\rangle$, norm square of $|u\rangle$ which is here and the norm of this vector which is norm of $|v_1\rangle$ times $|\langle v_1|v_2\rangle|^2$ over $|v_1|^4$. Now, this is the positive quantity so we can say, $|v_2|^2$ is always greater than or equal to $(|v_1|^2/|v_1|^4)|\langle v_1|v_2\rangle|^2$. This implies that $|v_2|^2|v_1|^2$ is greater than or equal to $|\langle v_1|v_2\rangle|^2$ or $|v_2||v_1|$ is greater than or equal to $\langle v_1|v_2\rangle$.

(Refer slide time: 32:12)



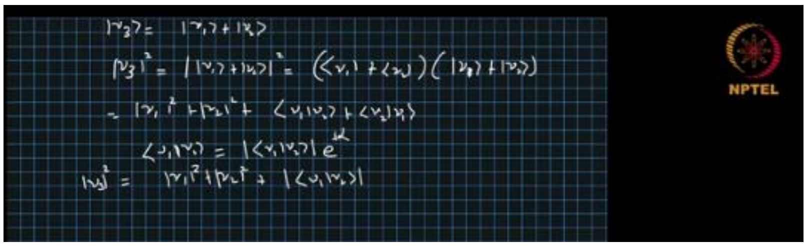
(Refer slide time: 32:43)



What we have achieved here is that the inner product of two vectors is always less than the product of their lengths. This is what we are saying. So, if you have two vectors, $|v_1\rangle$ and $|v_2\rangle$, and you take the product of their length, that will always be greater than the inner product of the two vectors, okay. So now, what we can, we will be using this relation in the following way, take $|v_3\rangle$, which let me remind you which is $|v_1\rangle + |v_2\rangle$, then $|v_3|^2$ is $\left||v_1\rangle + |v_1\rangle\right|^2$, which is $(|v_1\rangle + |v_2\rangle)(|v_1\rangle + |v_2\rangle)$ and that is equal to $|v_1|^2 + |v_2|^2 + \langle v_1|v_2\rangle + \langle v_2|v_1\rangle$.

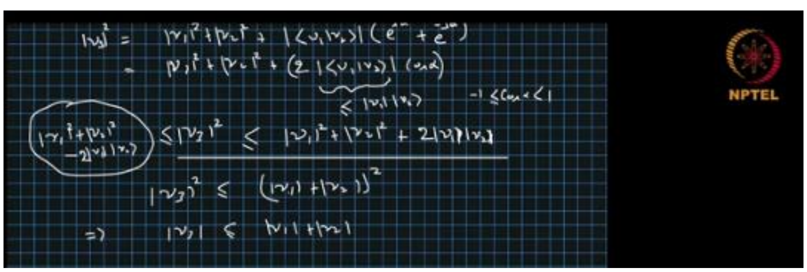
So, we have these four terms when we find the norm of the sum of $|v_1\rangle, |v_2\rangle$. Now, $\langle v_1|v_2\rangle$ is a complex number, if our field is complex. So, we can write it as $|\langle v_1|v_2\rangle|$ times the phase factor. So, it's a polar decomposition of a complex number. Any complex number can be written as a real positive number times a phase factor.

(Refer slide time: 34:37)



So, $e^{i\alpha}$ is our phase. So, $|v_3|^2$ can be written as $|v_1|^2 + |v_2|^2 + |\langle v_1|v_2\rangle|$. Just let me remind you that $\langle v_1|v_2\rangle$ is same as $\langle v_2|v_1\rangle$ and we have $e^{i\alpha}$ and the other one has to be complex conjugate. So, it is $e^{-i\alpha}$, which is $|v_1|^2 + |v_2|^2 + 2|\langle v_1|v_2\rangle|\ cos\alpha$. The $e^{i\alpha} + e^{-i\alpha}$ is $2\ cos\alpha$. Now, just remember our earlier inequality that this factor is always less than or equal to $|v_1||v_2|$, but the whole factor, we still don't know whether it will be less or more because $cos\alpha$ can take value between -1 and 1.

(Refer slide time: 37:36)



So, it means we can say that $|v_3|^2$ is always greater than $|v_1|^2 + |v_2|^2 - 2|v_1||v_2|$. And it is always less than or equal to $|v_1|^2 + |v_2|^2 + 2|v_1||v_2|$. What we have done here is, if $cos\alpha$ is, maximum value of this is +1, if we have, we take the maximum value of it that is plus one, then replacing $|\langle v_1|v_2\rangle|$ with $|v_1||v_2|$, then this quantity is always greater than $|v_3|^2$ and if $cos\alpha$ is negative and the maximum negative is -1, then, and then we replace $|\langle v_1|v_2\rangle|$ with $|v_1||v_2|$, then we get the lower bound on $|v_3|$. So, we get $|v_3|^2$ between this quantity, which is $|v_1|^2 + |v_2|^2 - 2|v_1||v_2|$ and $|v_1|^2 + |v_2|^2 + 2|v_1||v_2|$.

We are, for the time being only interested in this part. So, we get $|v_3|^2$ less than or equal to $(|v_1| + |v_2|)^2$. Hence, we get $|v_3|$ less than or equal to$|v_1 + v_2|$, and this is what was our intention to achieve.