

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high-quality educational resources for free. To make a donation or to view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at [ocw.mit.edu](https://ocw.mit.edu).

**PROFESSOR:**

Last time we talked about the spin operator pointing in some particular direction. There were questions. In fact, there was a useful question that I think I want to begin the lecture by going back to it. And this, you received an email from me. The notes have an extra section added to it that is stuff that I didn't do in class last time, but I was told in fact some of the recitation instructors did discuss this matter And I'm going to say a few words about it.

Now, I do expect you to read the notes. So things that you will need for the homework, all the material that is in the notes is material that I kind of assume you're familiar with. And you've read it and understood it.

And I probably don't cover all what is in the notes, especially examples or some things don't go into so much detail. But the notes should really be helping you understand things well.

So the remark I want to make is that-- there was a question last time that better that we think about it more deliberately in which we saw there that Pauli matrices,  $\sigma_1^2$  was equal to  $\sigma_2^2$  equal to  $2\sigma_3^2$  was equal to 1.

Well, that, indeed, tells you something important about the eigenvalues of this matrices. And it's a general fact. If you have some matrix  $M$  that satisfies an equation. Now, let me write an equation.

The matrix  $M^2$  plus  $\alpha M$  plus  $\beta$  times the identity is equal to 0. This is a matrix equation. It takes the whole matrix, square it, add  $\alpha$  times the matrix, and then  $\beta$  times the identity matrix is equal to 0.

Suppose you discover that such an equation holds for that matrix  $M$ . Then, suppose you are also asked to find eigenvalues of this matrix  $M$ . So suppose there is a vector-- that is, an eigenvector with eigenvalue  $\lambda$ . That's what having an eigenvector with eigenvalue  $\lambda$  means. And you're supposed to calculate these values of  $\lambda$ .

So what you do here is let this equation, this matrix on the left, act on the vector  $v$ . So you have  $M^2 + \alpha M + \beta I$  act on  $v$ . Since the matrix is 0, it should be 0. And now you come and say, well, let's see.  $\beta I$  on  $v$ . Well, that's just  $\beta v$ , the vector  $v$ .

$\alpha M$  on  $v$ , but  $M$  on  $v$  is  $\lambda v$ . So this is  $\alpha \lambda v$ . And  $M^2$  on  $v$ , as you can imagine, you act with another  $M$  here. Then you go to this side. You get  $\lambda Mv$ , which is, again, another  $\lambda$  times  $v$ . So  $M^2$  on  $v$  is  $\lambda^2 v$ . If acts two times on  $v$ .

Therefore, this is 0. And here you have, for example, that  $\lambda^2 + \alpha \lambda + \beta$  on  $v$  is equal to 0. Well,  $v$  cannot be 0. Any eigenvector-- by definition, eigenvectors are not 0 vectors. You can have 0 eigenvalues but not 0 eigenvectors. That doesn't exist. An eigenvector that is 0 is a crazy thing because this would be 0, and then it would be-- the eigenvalue would not be determined. It just makes no sense. So  $v$  is different from 0.

So you see that  $\lambda^2 + \alpha \lambda + \beta = 0$ . And the eigenvalues, any eigenvalue of this matrix, must satisfy this equation. So the eigenvalues of  $\sigma_1$ , you have  $\sigma_1^2$ , for example, is equal to 1. So the eigenvalues, any  $\lambda^2$  must be equal to 1, the number 1.

And therefore, the eigenvalues of  $\sigma_1$  are possibly plus or minus 1. We don't know yet. Could be two 1's, 2 minus 1's, one 1 and one minus 1. But there's another nice thing, the trace of  $\sigma_1$ . We'll study more the trace, don't worry. If you are not that familiar with it, it will become more familiar soon.

The trace of  $\sigma_1$  or any matrix is the sum of elements in the diagonal.  $\sigma_1$ ,

if you remember, was of this form. Therefore, the trace is 0. And in fact, the traces of any of the Pauli matrices are 0.

Another little theorem of linear algebra shows that the trace of a matrix is equal to the sum of eigenvalues. So whatever two eigenvalues  $\sigma_1$  has, they must add up to 0. Because the trace is 0 and it's equal to the sum of eigenvalues.

And therefore, if the eigenvalues can only be plus or minus 1, you have the result that one eigenvalue must be plus 1. The other eigenvalue must be minus 1, is the only way you can get that to work.

So two  $\sigma_1$  eigenvalues of  $\sigma_1$  are plus 1 and minus 1. Those are the two eigenvalues. So in that section as well, there's some discussion about properties of the Pauli matrices.

And two basic properties of Pauli matrices are the following. Remember that the spin matrices, the spin operators, are  $\hbar/2$  times the Pauli matrices. And the spin operators had the algebra for angular momentum. So from the algebra of angular momentum that says that  $S_i S_j$  is equal to  $i \hbar \epsilon_{ijk} S_k$ , you deduce after plugging this that  $\sigma_i \sigma_j$  is  $2i \epsilon_{ijk} \sigma_k$ .

Moreover, there's another nice property of the Pauli matrices having to deal with anticommutators. If you do experimentally try multiplying Pauli matrices,  $\sigma_1$  and  $\sigma_2$ , you will find out that if you compare it with  $\sigma_2 \sigma_1$ , it's different. Of course, it's not the same. These matrices don't commute. But they actually-- while they fail to commute, they still fail to commute in a nice way. Actually, these are minus each other. So in fact,  $\sigma_1 \sigma_2 + \sigma_2 \sigma_1$  is equal to 0. And by this, we mean that they anticommute. And we have a brief way of calling this.

When this sign was a minus, it was called the commutator. When this is a plus, it's called an anticommutator. So the anticommutator of  $\sigma_1$  with  $\sigma_2$  is equal to 0. Anticommutator defined in general by A, B. Two operators is  $AB + BA$ .

And as you will read in the notes, a little more analysis shows that, in fact, the

anticommutator of  $\sigma_i$  and  $\sigma_j$  has a nice formula, which is  $2\delta_{ij}$  times the unit matrix, the 2 by 2 unit matrix.

With this result, you get a general formula. Any product of two operators,  $AB$ , you can write as  $1/2$  of the anticommutator plus  $1/2$  of the commutator plus  $1/2$  of the anticommutator.

Expand it out, that right-hand side, and you will see quite quickly this is true for any two operators. This has  $AB$  minus  $BA$  and this has  $AB$  plus  $BA$ . The  $BA$  term cancels and the  $AB$  terms are  $[2AB]$ . So  $\sigma_i \sigma_j$  would be equal to  $1/2$ . And then they put down the anticommutator first. So you get  $\delta_{ij}$  times the identity, which is  $1/2$  of the anticommutator plus  $1/2$  of the commutator, which is  $i \epsilon_{ijk} \sigma_k$ . It's a very useful formula.

In order to make those formulas look neater, we invent a notation in which we think of  $\sigma$  as a triplet--  $\sigma_1$ ,  $\sigma_2$ , and  $\sigma_3$ . And then we have vectors, like  $a$ -- normal vectors, components  $a_1$ ,  $a_2$ ,  $a_3$ . And then we have a dot  $\sigma$  must be defined.

Well, there's an obvious definition of what this should mean, but it's not something you're accustomed to. And one should pause before saying this. You're having a normal vector, a triplet of numbers, multiplied by a triplet of matrices, or a triplet of operators. Since numbers commute with matrices, the order in which you write this doesn't matter. But this is defined to be  $a_1 \sigma_1$  plus  $a_2 \sigma_2$  plus  $a_3 \sigma_3$ .

This can be written as  $a_i \sigma_i$  with our repeated index convention that you sum over the possibilities. So here is what you're supposed to do here to maybe interpret this equation nicely. You multiply this equation  $n$  by  $a_j$ .

Now, these are numbers. These are matrices. I better not change this order, but I can certainly, by multiplying that way, I have  $a_i \sigma_i a_j \sigma_j$  equals  $2 a_i a_j \delta_{ij}$  plus  $i \epsilon_{ijk} a_i a_j \sigma_k$ . Now, what?

Well, write it in terms of things that look neat.  $a \cdot \sigma$ , that's a matrix. This

whole thing is a matrix multiplied by the matrix  $b \cdot \sigma$  gives you--

Well,  $a_i b_j \delta_{ij}$ , this  $\delta_{ij}$  forces  $j$  to become  $i$ . In other words, you can replace these two terms by just  $b_i$ . And then you have  $a_i b_i$ . So this is twice. I don't know why I have a 2. No 2. There was no 2 there, sorry. So what do we get here?

We get  $a \cdot b$ , the dot product. This is a normal dot product. This is just a number times 1 plus  $i$ . Now, what is this thing?

You should try to remember how the epsilon tensor can be used to do cross products. This, there's just one free index, the index  $k$ . So this must be some sort of vector. And in fact, if you try the definition of epsilon and look in detail what this is, you will find that this is nothing but the  $k$  component of  $a \cdot b$ . The  $k$ -- so I'll write it here.

This is a cross  $b$  sub  $k$ . But now you have a cross  $b$  sub  $k$  times  $\sigma_k$ . So this is the same as a cross  $b \cdot \sigma$ . And here you got a pretty nice equation for Pauli matrices. It expresses the general product of Pauli matrices in somewhat geometric terms.

So if you take, for example here, an operator. No. If you take, for example,  $a$  equals  $b$  equal to a unit vector, then what do we get?

You get  $n \cdot \sigma$  squared. And here you have the dot product of  $n$  with  $n$ , which is 1. So this is 1. And the cross product of two equal vectors, of course, is 0 so you get this, which is nice. Why is this useful?

It's because with this identity, you can understand better the operator  $S_{\hat{n}}$  that we introduced last time, which was  $n \cdot \text{spin triplet}$ . So  $n_x, n_y, n_z, s_x, s_y, s_z$ . So what is this?

This is  $\frac{\hbar}{2} n \cdot \sigma$ . And let's square this. So  $S_{\hat{n}}$  vector squared. This matrix squared would be  $\frac{\hbar^2}{4} n \cdot \sigma$  squared, which is 1. And  $\sigma$  squared is 1. Therefore, this spin operator along the  $n$  direction squares to  $\frac{\hbar^2}{4}$  times 1.

Now, the trace of this  $S_n$  operator is also 0. Why?

Because the trace means that you're going to sum the elements in the diagonal. Well, you have a sum of matrices here. And therefore, you will have to sum the diagonals of each. But each of the sigmas has 0 trace. We wrote it there. Trace of sigma 1 is 0. All the Pauli matrices have 0 trace, so this has 0 trace. So you have these two relations.

And again, this tells you that the eigenvalues of this matrix can be plus minus  $\hbar/2$ . Because the eigenvalues satisfy the same equation as the matrix. Therefore, plus minus  $\hbar/2$ . And this one says that the eigenvalues add up to 0. So the eigenvalues of  $S_n$  vector are plus  $\hbar/2$  and minus  $\hbar/2$ .

We did that last time, but we do that by just taking that matrix and finding the eigenvalues. But this shows that its property is almost manifest. And this is fundamental for the interpretation of this operator. Why?

Well, we saw that if  $n$  points along the z-direction, it becomes the operator  $s_z$ . If it points about the x-direction, it becomes the operator  $s_x$ . If it points along y, it becomes  $s_y$ . But in an arbitrary direction, it's a funny thing. But it still has the key property.

If you measured the spin along an arbitrary direction, you should find only plus  $\hbar/2$  or minus  $\hbar/2$ . Because after all, the universe is isotropic. It doesn't depend on direction. So a spin one-half particle. If you find out that whenever you measure the z component, it's either plus minus  $\hbar/2$ . Well, when you measure any direction, it should be plus minus  $\hbar/2$ .

And this shows that this operator has those eigenvalues. And therefore, it makes sense that this is the operator that measures spins in an arbitrary direction.

There's a little more of an aside in there, in the notes about something that will be useful and fun to do. And it corresponds to the case in which you have two triplets of

operators--  $x_1, x_2, x_3$ . These are operators now. And  $y$  equal  $y_1, y_2, y_3$ . Two triplets of operators.

So you define the dot product of these two triplets as  $x_i y_i$  summed. That's the definition.

Now, the dot product of two triplets of operators defined that way may not commute. Because the operators  $x$  and  $y$  may not commute. So this new dot product of both phase operators is not commutative-- probably. It may happen that these operators commute, in which case  $x \text{ dot } y$  is equal to  $y \text{ dot } x$ .

Similarly, you can define the cross product of these two things. And the  $k$ -th component is  $\epsilon_{ijk} x_i y_j$  like this. Just like you would define it for two number vectors. Now, what do you know about the cross product in general?

It's anti-symmetric.  $A \text{ cross } B$  is equal to minus  $B \text{ cross } A$ . But this one won't be because the operators  $x$  and  $y$  may not commute. Even  $x \text{ cross } x$  may be nonzero. So one thing I will ask you to compute in the homework is not a long calculation. It's three lines. But what is  $S \text{ cross } S$  equal to? Question there?

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** Yes, it's the sum [INAUDIBLE]. Just in the same way that here you're summing over  $i$ 's and  $j$ 's to produce the cross product. So whenever an index is repeated, we'll assume it's summed. And when it is not summed, I will put to the right, not summed explicitly-- the words-. Because in some occasions, it matters.

So how much is this? It will involve  $i, \hbar$ , and something. And you will try to find out what this is. It's a cute thing. All right, any other questions? More questions? Nope. OK.

So now, finally, we get to that part of the course that has to do with linear algebra. And I'm going to do an experiment. I'm going to do it differently than I did it in the previous years.

There is this nice book. It's here. I don't know if you can read from that far, but it has

a pretty-- you might almost say an arrogant title. It says, *Linear Algebra Done Right* by Sheldon Axler. This is the book, actually, MIT's course 18.700 of linear algebra uses.

And when you first get the book that looks like that, you read it and open-- I'm going to show you that this is not that well done. But actually, I think it's actually true. The title is not a lie. It's really done right.

I actually wish I had learned linear algebra this way. It may be a little difficult if you've never done any linear algebra. You don't know what the matrix is-- I don't think that's the case anybody here. A determinant, or eigenvalue. If you never heard any of those words, this might be a little hard. But if you've heard those words and you've had a little linear algebra, this is quite nice.

Now, this book has also a small problem. Unless you study it seriously, it's not all that easy to grab results that you need from it. You have to study it. So I don't know if it might help you or not during this semester. It may.

It's not necessary to get it. Absolutely not. But it is quite lovely. And the emphasis is quite interesting. It really begins from very basic things and logically develops everything and asks at every point the right questions. It's quite nice. So what I'm going to do is-- inspired by that, I want to introduce some of the linear algebra little by little. And I don't know very well how this will go. Maybe there's too much detail. Maybe it's a lot of detail, but not enough so it's not all that great. I don't know, you will have to tell me. But we'll try to get some ideas clear.

And the reason I want to get some ideas clear is that good books on this subject allow you to understand how much structure you have to put in a vector space to define certain things. And unless you do this carefully, you probably miss some of the basic things. Like many physicists don't quite realize that talking about the matrix representation, you don't need brass and [INAUDIBLE] to talk about the matrix representation of an operator.

At first sight, it seems like you'd need it, but you actually don't. Then, the differences



between a complex and a vector space-- complex and a real vector space become much clearer if you take your time to understand it. They are very different. And in a sense, complex vector spaces are more powerful, more elegant, have stronger results.

So anyway, it's enough of an introduction. Let's see how we do. And let's just begin there for our story. So we begin with vector spaces and dimensionality. Yes.

**AUDIENCE:** Quick question. The length between the trace of matrix equals 0 and [INAUDIBLE] is proportional to the identity. One is the product of the eigenvalues is 1 and the other one was the sum is equal to 0. Are those two statements related causally, or are they just separate statements [INAUDIBLE]?

**PROFESSOR:** OK, the question is, what is the relation between these two statements? Those are separate observations. One does not imply the other. You can have matrices that square to the identity, like the identity itself, and don't have 0 trace. So these are separate properties.

This tells us that the eigenvalue squared are  $\hbar^2$ . And this one tells me that  $\lambda_1 + \lambda_2$ -- there are two eigenvalues-- are 0. So from here, you deduce that the eigenvalues could be plus minus  $\hbar$ . And in fact, have to be plus minus  $\hbar$ .

All right, so let's talk about vector spaces and dimensionality. Spaces and dimensionality. So why do we care about this?

Because the end result of our discussion is that the states of a physical system are vectors in a complex vector space. That's, in a sense, the result we're going to get.

Observables, moreover, are linear operators on those vector spaces. So we need to understand what are complex vector spaces, what linear operators on them mean.

So as I said, complex vector spaces have subtle properties that make them different from real vector spaces and we want to appreciate that. In a vector space, what do you have?

You have vectors and you have numbers. So the two things must exist. The numbers could be the real numbers, in which case we're talking about the real vector space. And the numbers could be complex numbers, in which case we're talking about the complex vector space. We don't say the vectors are real, or complex, or imaginary. We just say there are vectors and there are numbers.

Now, the vectors can be added and the numbers can be multiplied by vectors to give vectors. That's basically what is happening.

Now, these numbers can be real or complex. And the numbers-- so there are vectors and numbers. And we will focus on just either real numbers or complex numbers, but either one. So these sets of numbers form what is called in mathematics a field. So I will not define the field. But a field-- use the letter  $F$  for field. And our results. I will state results whenever-- it doesn't matter whether it's real or complex, I may use the letter  $F$  to say the numbers are in  $F$ . And you say real or complex.

What is a vector space? So the vector space,  $V$ . Vector space,  $V$ , is a set of vectors with an operation called addition-- and we represent it as plus-- that assigns a vector  $u$  plus  $v$  in the vector space when  $u$  and  $v$  belong to the vector space.

So for any  $u$  and  $v$  in the vector space, there's a rule called addition that assigns another vector. This also means that this space is closed under addition. That is, you cannot get out of the vector space by adding vectors. The vector space must contain a set that is consistent in that you can add vectors and you're always there.

And there's a multiplication. And a scalar multiplication by elements of the numbers of  $F$  such that  $a$ , which is a number, times  $v$  belongs to the vector space when  $a$  belongs to the numbers and  $v$  belongs to the vectors.

So every time you have a vector, you can multiply by those numbers and the result of that multiplication is another vector. So we say the space is also closed under multiplication.

Now, these properties exist, but they must-- these operations exist, but they must satisfy the following properties. So the definition is not really over. These operations satisfy--

1.  $u + v$  is equal to  $v + u$ . The order doesn't matter how you sum vectors. And here,  $u$  and  $v$  in  $V$ .

2. Associative. So  $u + v + w$  is equal to  $u + (v + w)$ . Moreover, two numbers  $a$  times  $b$  times  $v$  is the same as  $(a \times b)$  times  $v$ . You can add with the first number on the vector and you add with the second.

3. There is an additive identity. And that is what?

It's a vector  $0$  belonging to the vector space. I could write an arrow. But actually, for some reason they just don't like to write it because they say it's always ambiguous whether you're talking about the  $0$  number or the  $0$  vector. We do have that problem also in the notation in quantum mechanics. But here it is, here is a  $0$  vector such that  $0$  plus any vector  $v$  is equal to  $v$ .

4. Well, in the field, in the set of numbers, there's the number  $1$ , which multiplied by any other number keeps that number. So the number  $1$  that belongs to the field satisfies that  $1$  times any vector is equal to the vector. So we declare that that number multiplied by other numbers is an identity. [INAUDIBLE] identity also multiplying vectors. Yes, there was a question.

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** There is an additive identity. Additive identity, the  $0$  vector.

Finally, distributive laws. No. One second. One, two, three-- the zero vector.

Oh, actually in my list I put them in different orders in the notes, but never mind.

5. There's an additive inverse in the vector space. So for each  $v$  belonging to the vector space, there is a  $u$  belonging to the vector space such that  $v + u$  is equal to  $0$ . So additive identity you can find for every element its opposite vector. It always

can be found.

And last is this [INAUDIBLE] which says that  $a$  times  $u$  plus  $v$  is equal to  $au$  plus  $av$ , and  $a$  plus  $b$  on  $v$  is equal to  $av$  plus  $bv$ . And  $a$ 's and  $b$ 's belong to the numbers.  $a$  and  $b$ 's belong to the field. And  $u$  and  $v$  belong to the vector space. OK.

It's a little disconcerting. There's a lot of things. But actually, they are quite minimal. It's well done, this definition. They're all kind of things that you know that follow quite immediately by little proofs. You will see more in the notes, but let me just say briefly a few of them.

So here is the additive identity, the vector  $0$ . It's easy to prove that this vector  $0$  is unique. If you find another  $0$  prime that also satisfies this property,  $0$  is equal to  $0$  prime. So it's unique.

You can also show that  $0$  times any vector is equal to  $0$ . And here, this  $0$  belongs to the field and this  $0$  belongs to the vector space. So the  $0$ -- you had to postulate that the  $1$  in the field does the right thing, but you don't need to postulate that  $0$ , the number  $0$ , multiplied by a vector is  $0$ . You can prove that. And these are not difficult to prove. All of them are one-line exercises. They're done in that book. You can look at them.

Moreover, another one.  $a$  any number times the  $0$  vector is equal to the  $0$  vector. So in this case, those both are vectors. That's also another property. So the  $0$  vector and the  $0$  number really do the right thing.

Then, another property, the additive inverse. This is sort of interesting. So the additive inverse, you can prove it's unique. So the additive inverse is unique. And it's called-- for  $v$ , it's called minus  $v$ , just a name. And actually, you can prove it's equal to the number minus  $1$  times the vector.

Might sound totally trivial but try to prove them. They're all simple, but they're not trivial, all these things. So you call it minus  $v$ , but it's actually-- this is a proof.

OK. So examples. Let's do a few examples. I'll have five examples that we're going

to use.

So I think the main thing for a physicist that I remember being confused about is the statement that there's no characterization that the vectors are real or complex. The vectors are the vectors and you multiply by a real or complex numbers. So I will have one example that makes that very dramatic. As dramatic as it can be.

So one example of vector spaces, the set of  $N$  component vectors. So here it is,  $a_1$ ,  $a_2$ , up to  $a_n$ . For example, with capital  $N$ . With  $a_i$  belongs to the real and  $i$  going from 1 up to  $N$  is a vector space over  $\mathbb{R}$ , the real numbers. So people use that terminology, a vector space over the kind of numbers. You could call it also a real vector space, that would be the same.

You see, these components are real. And you have to think for a second if you believe all of them are true or how would you do it.

Well, if I would be really precise, I would have to tell you a lot of things that you would find boring. That, for example, you have this vector and you add a set of  $b$ 's. Well, you add the components. That's the definition of plus. And what's the definition of multiplying by a number?

Well, if a number is multiplied by this vector, it goes in and multiplies everybody. Those are implicit, or you can fill-in the details. But if you define them that way, it will satisfy all the properties. What is the 0 vector? It must be the one with all entries 0. What is the additive inverse?

Well, change the sign of all these things. So it's kind of obvious that this satisfies everything, if you understand how the sum and the multiplication goes.

Another one, it's kind of similar. 2. The set of  $M$  cross  $N$  matrices with complex entries. Complex entries. So here you have it,  $a_{11}$ ,  $a_{12}$ ,  $a_{1N}$ . And here it goes up to  $a_{M1}$ ,  $a_{M2}$ ,  $a_{MN}$ . With all the  $a_{ij}$ 's belonging to the complex numbers, then-- I'll erase here. Then you have that this is a complex vector space. Is a complex vector space.

How do you multiply by a number? You multiply a number times every entry of the matrices.

How do sum two matrices? They have the same size, so you sum each element the way it should be. And that should be a vector space. Here is an example that is, perhaps, a little more surprising.

So the space of 2 by 2 Hermitian matrices is a real vector space.

You see, this can be easily thought [INAUDIBLE] naturally thought as a real vector space. This is a little surprising because Hermitian matrices have  $i$ 's. You remember the most general Hermitian matrix was of the form-- well,  $a + ib$ ,  $a - ib$ , with all these numbers  $a, b, c, d$  in real. But they're complex numbers. Why is this naturally a real vector space?

The problem is that if you multiply by a number, it should still be a Hermitian matrix in order for it to be a vector space. It should be in the vector. But if you multiply by a real number, there's no problem. The matrix remains Hermitian. You multiplied by a complex number, you use the Hermiticity. But an  $i$  somewhere here for all the factors and it will not be Hermitian. So this is why it's a real vector space.

Multiplication by real numbers preserves Hermiticity.

So that's surprising. So again, illustrates that nobody would say this is a real vector. But it really should be thought as a vector over real numbers. Vector space over real numbers.

Two more examples. And they are kind of interesting. So the next example is the set of polynomials as vector space. So that, again, is sort of a very imaginative thing. The set of polynomials  $p$  of  $z$ .

Here,  $z$  belongs to some field and  $p$  of  $z$ , which is a function of  $z$ , also belongs to the same field. And each polynomial has coefficient. So any  $p$  of  $z$  is  $a_0$  plus  $a_1 z$  plus  $a_2 z^2$  plus-- up to some  $a_n z^n$ . A polynomial is supposed to end That's pretty important about polynomials. So the dots don't go up forever.

So here it is, the  $a_i$ 's also belong to the field. So looked at this polynomials. We have the letter  $z$  and they have these coefficients which are numbers. So a real polynomial-- you know  $2 + x + x^2$ . So you have your real numbers times this general variable that it's also supposed to be real. So you could have it real. You could have it complex. So that's a polynomial. How is that a vector space?

Well, it's a vector space-- the space  $P$  of  $F$  of those polynomials-- of all polynomials is a vector space over  $F$ . And why is that?

Well, you can take-- again, there's some implicit definitions. How do you sum polynomials?

Well, you sum the independent coefficients. You just sum them and factor out. So there's an obvious definition of sum. How do you multiply a polynomial by a number?

Obvious definition, you multiply everything by a number. If you sum polynomials, you get polynomials. Given a polynomial, there is a negative polynomial that adds up to 0. There's a 0 when all the coefficients is 0. And it has all the nice properties.

Now, this example is more nontrivial because you would think, as opposed to the previous examples, that this is probably infinite dimensional because it has the linear polynomial, the quadratic, the cubic, the quartic, the quintic, all of them together. And yes, we'll see that in a second. So set of polynomials.

5. Another example, 5. The set  $F$  infinity of infinite sequences. Sequences  $x_1, x_2, \dots$  infinite sequences where the  $x_i$ 's are in the field. So you've got an infinite sequence and you want to add another infinite sequence.

Well, you add the first element, the second elements. It's like an infinite column vector. Sometimes mathematicians like to write column vectors like that because it's practical. It saves space on a page. The vertical one, you start writing and the pages grow very fast.

So here's an infinite sequence. And think of it as a vertical one if you wish. And all

elements are here, but there are infinitely many in every sequence. And of course, the set of all infinite sequences is infinite. So this is a vector space over  $F$ . Again, because all the numbers are here, so it's a vector space over  $F$ .

And last example. Our last example is a familiar one in physics, is the set of complex functions in an interval. Set of complex functions on an interval  $x$  from 0 to  $L$ . So a set of complex functions  $f$  of  $x$  I could put here on an interval [INAUDIBLE]. So this is a complex vector space. Vector space.

The last three examples, probably you would agree that there are infinite dimensional, even though I've not defined what that means very precisely. But that's what we're going to try to understand now. We're supposed to understand the concept of dimensionality. So let's get to that concept now.

So in terms of dimensionality, to build this idea you need a definition. You need to know the term subspace of a vector space. What is a subspace of a vector space?

A subspace of a vector space is a subset of the vector space that is still a vector space. So that's why it's called subspace. It's different from subset. So a subspace of  $V$  is a subset of  $V$  that is a vector space.

So in particular, it must contain the vector 0 because any vector space contains the vector 0. One of the ways you sometimes want to understand the vector space is by representing it as a sum of smaller vector spaces. And we will do that when we consider, for example, angular momentum in detail. So you want to write a vector space as a sum of subspaces. So what is that called?

It's called a direct sum. So if you can write-- here is the equation. You say  $V$  is equal to  $u_1$  direct sum with  $u_2$  direct sum with  $u_3$  direct sum with  $u_m$ . When we say this, we mean the following.

That the  $u_i$ 's are subspaces of  $V$ . And any  $V$  in the vector space can be written uniquely as  $a_1 u_1$  plus  $a_2 u_2$  plus  $a_n u_n$  with  $u_i$  [INAUDIBLE] capital  $U_i$ . So let me review what we just said.



So you have a vector space and you want to decompose it in sort of basic ingredients. This is called a direct sum.  $V$  is a direct sum of subspaces. Direct sum. And the  $U_i$ 's are subspaces of  $V$ . But what must happen for this to be true is that once you take any vector here, you can write it as a sum of a vector here, a vector here, a vector here, a vector everywhere. And it must be done uniquely.

If you can do this in more than one way, this is not a direct sum. These subspaces kind of overlap. They're not doing the decomposition in a minimal way. Yes.

**AUDIENCE:** Does the expression of  $V$  have to be a linear combination of the vectors of the  $U$ , or just sums of the  $U$  sub  $i$ 's?

**PROFESSOR:** It's some linear combination. Look, the interpretation, for example,  $\mathbb{R}^2$ . The normal vector space  $\mathbb{R}^2$ . You have an intuition quite clearly that any vector here is a unique sum of this component along this subspace and this component along this subspace. So it's a trivial example, but the vector space  $\mathbb{R}^2$  has a vector subspace  $\mathbb{R}^1$  here and a vector subspace  $\mathbb{R}^1$ . Any vector in  $\mathbb{R}^2$  is uniquely written as a sum of these two vectors. That means that  $\mathbb{R}^2$  is really  $\mathbb{R}^1$  plus  $\mathbb{R}^1$ . Yes.

**AUDIENCE:** [INAUDIBLE]. Is it redundant to say that that-- because  $a_1 u_1$  is also in big  $U$  sub 1.

**PROFESSOR:** Oh. Oh, yes. You're right. No, I'm sorry. I shouldn't write those. I'm sorry. That's absolutely right. If I had that in my notes, it was a mistake. Thank you. That was very good. Did I have that in my notes? No, I had it as you said it. True. So can be written uniquely as a vector in first, a vector in the second. And the  $a$ 's are absolutely not necessary. OK. So let's go ahead then and say the following things.

So here we're going to try to get to the concept of dimensionality in a precise way. Yes.

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** Right, the last one is  $m$ . Thank you. All right.

The concept of dimensionality of a vector space is something that you intuitively understand. It's sort of how many linearly independent vectors you need to describe

the whole set of vectors. So that is the number you're trying to get to. I'll follow it up in a slightly rigorous way to be able to do infinite dimensional space as well.

So we will consider something called a list of vectors. List of vectors. And that will be something like  $v_1, v_2$  vectors in a vector space up to  $v_n$ .

Any list of vectors has finite length. So we don't accept infinite lists by definition. You can ask, once you have a list of vectors, what is the vector subspace spanned by this list? How much do you reach with that list?

So we call it the span of the list. The span of the list,  $v_n$ . And it's the set of all linear combinations  $a_1 v_1$  plus  $a_2 v_2$  plus  $a_n v_n$  for  $a_i$  in the field. So the span of the list is all possible products of your vectors on the list are-- and put like that.

So if we say that the list spans a vector space, if the span of the list is the vector space. So that's natural language. We say, OK, this list spans the vector space. Why?

Because if you produce the span of the list, it fills a vector space. OK, so I could say it that way. So here is the definition,  $V$  is finite dimensional if it's spanned by some list. If  $V$  is spanned by some list. So why is that?

Because if the list is-- a definition, finite dimensional. If it's spanned by some list. If you got your list, by definition it's finite length. And with some set of vectors, you span everything.

And moreover, it's infinite dimensional if it's not finite dimensional. It's kind of silly, but infinite-- a space  $V$  is infinite dimensional if it is not finite dimensional. Which is to say that there is no list that spans the space.

So for example, this definition is tailored in a nice way. Like let's think of the polynomials. And we want to see if it's finite dimensional or infinite dimensional. So you claim it's finite dimensional. Let's see if it's finite dimensional.

So we make a list of polynomials. The list must have some length, at least, that

spans it. You put all these 730 polynomials that you think span the list, span the space, in this list.

Now, if you look at the list, it's 720. You can check one by one until you find what is the one of highest order, the polynomial of highest degree.

But if the highest degree is say,  $z$  to the 1 million, then any polynomial that has a  $z$  to the 2 million cannot be spanned by this one. So there's no finite list that can span this, so this set-- the example in 4 is infinite dimensional for sure. Example 4 is infinite dimensional.

Well, example one is finite dimensional. You can see that because we can produce a list that spans the space. So look at the example 1.

It's there. Well, what would be the list?

The list would be-- list. You would put a vector  $e_1, e_2, \dots, e_n$ . And the vector  $e_1$  would be  $1, 0, 0, 0, 0$ . The vector  $e_2$  would be  $0, 1, 0, 0, 0$ . And go on like that. So you put 1's and 0's. And you have  $n$  of them. And certainly, the most general one is  $a_1$  times  $e_1$   $a_2$  times  $e_2$ . And you got the list. So example 1 is finite dimensional.

A list of vectors is linearly independent. A list is linearly independent if a list  $v_1$  up to  $v_n$  is linearly independent, if  $a_1 v_1$  plus  $a_2 v_2$  plus  $a_n v_n$  is equal to 0 has the unique solution  $a_1$  equal  $a_2$  equal all of them equal 0. So that is to mean that whenever this list satisfies this property-- if you want to represent the vector 0 with this list, you must set all of them equal to 0, all the coefficients. That's clear as well in this example.

If you want to represent the 0 vector, you must have 0 component against the basis vector  $x$  and basis vector  $y$ . So the list of this vector and this vector is linearly independent because the 0 vector must have 0 numbers multiplying each of them. So finally, we define what is a basis.

A basis of  $V$  is a list of vectors in  $V$  that spans  $V$  and is linearly independent. So what is a basis?

Well, you should have enough vectors to represent every vector. So it must span  $V$ . And what else should it have?

It shouldn't have extra vectors that you don't need. It should be minimal. It should be all linearly independent. You shouldn't have added more stuff to it. So any finite dimensional vector space has a basis. It's easy to do it.

There's another thing that one can prove. It may look kind of obvious, but it requires a small proof that if you have-- the bases are not unique. It's something we're going to exploit all the time. One basis, another basis, a third basis. We're going to change basis all the time.

Well, the bases are not unique, but the length of the bases of a vector space is always the same. So the length of the list is-- a number is the same whatever base you choose. And that length is what is called the dimension of the vector space.

So the dimension of a vector space is the length of any bases of  $V$ . And therefore, it's a well-defined concept. Any base of a finite vector space has the same length, and the dimension is that number. So there was a question. Yes?

**AUDIENCE:** Is there any difference between bases [INAUDIBLE]?

**PROFESSOR:** No, absolutely not. You could have a basis, for example, of  $\mathbb{R}^2$ , which is this vector. The first and the second is this vector. And any vector is a linear superposition of these two vectors with some coefficients and it's unique. You can find the coefficients.

**AUDIENCE:** [INAUDIBLE].

**PROFESSOR:** Yes. But you see, here is exactly what I wanted to make clear. We're putting the vector space and we're putting the least possible structure. I didn't say how to take the inner product of two vectors. It's not a definition of a vector space. It's something we'll put later.

And then, we will be able to ask whether the basis is orthonormal or not. But the

basis exists. Even though you have no definition of an inner product, you can talk about basis without any confusion. You can also talk about the matrix representation of an operator. And you don't need an inner product, which is sometimes very unclear.

You can talk about the trace of an operator and you don't need an inner product. You can talk about eigenvectors and eigenvalues and you don't need an inner product.

The only thing you need the inner product is to get numbers. And we'll use them to use [INAUDIBLE] to get numbers. But it can wait. It's better than you see all that you can do without introducing more things, and then introduce them. So let me explain a little more this concept.

We were talking about this base, this vector space  $V$ , for example. And we produced a list that spans  $e_1, e_2, \dots, e_n$ . And those were these vectors.

Now, this list not only spans, but they are linearly independent. If you put  $a_1$  times this plus  $a_2$  times this and you set it all equal to 0. Well, each entry will be 0, and all the  $a$ 's are 0. So these  $e$ 's that you put here on that list is actually a basis.

Therefore, the length of that basis is the dimensionality. And this space has dimension  $N$ . You should be able to prove that this space has been dimension  $m$  times  $N$ . Now, let me do the Hermitian-- these matrices. And try to figure out the dimensionality of the space of Hermitian matrices.

So here they are. This is the most general Hermitian matrix. And I'm going to produce for you a list of four vectors. Vectors-- yes, they're matrices, but we call them vectors. So here is the list. The unit matrix, the first Pauli matrix, the second Pauli matrix, and the third Pauli matrix. All right, let's see how far do we get from there.

OK, this is a list of vectors in the vector space because all of them are Hermitian. Good. Do they span?

Well, you calculate the most general Hermitian matrix of this form. You just put arbitrary complex numbers and require that the matrix be equal to its matrix complex conjugate and transpose. So this is the most general one. Do I obtain this matrix from this one's?

Yes I just have to put  $1$  times  $c$  plus  $a$  times  $\sigma_1$  plus  $b$  times  $\sigma_2$  plus  $d$  times  $\sigma_3$ . So any Hermitian matrix can be obtained as the span of this list. Is this list linearly independent?

So I have to go here and set this equal to  $0$  and see if it sets to  $0$  all these coefficients. Well, it's the same thing as setting to  $0$  all this matrix.

Well, if  $c + d$  and  $c - d$  are  $0$ , then  $c$  and  $d$  are  $0$ . If this is  $0$ , it must be a  $0$  and  $b = 0$ , so all of them are  $0$ . So yes, it's linearly independent. It spans. Therefore, you've proven completely rigorously that this vector space is dimension  $4$ .

This vector space-- I will actually leave it as an exercise for you to show that this vector space is infinite dimensional. You say, of course, it's infinite dimensional. It has infinite sequences.

Well, you have to show that if you have a finite list of those infinite sequences, like  $300$  sequences, they span that. They cannot span that. So it takes a little work. It's interesting to think about it. I think you will enjoy trying to think about this stuff.

So that's our discussion of dimensionality. So this one is a little harder to make sure it's infinite dimensional. And this one is, yet, a bit harder than that one but it can also be done. This is infinite dimensional. And this is infinite dimensional.

In the last two minutes, I want to tell you a little bit-- one definition and let you go with that, is the definition of a linear operator.

So here is one thing. So you can be more general, and we won't be that general. But when you talk about linear maps, you have one vector space and another vector space,  $v$  and  $w$ . This is a vector space and this is a vector space.

And in general, a map from here is sometimes called, if it satisfies the property, a

linear map. And the key thing is that in all generality, these two vector spaces may not have the same dimension. It might be one vector space and another very different vector space. You go from one to the other.

Now, when you have a vector space  $v$  and you map to the same vector space, this is also a linear map, but this is called an operator or a linear operator. And what is a linear operator therefore?

A linear operator is a function  $T$ . Let's call the linear operator  $T$ . It takes  $v$  to  $v$ . In which way?

Well,  $T$  acting  $u$  plus  $v$ , on the sum of vectors, is  $Tu$  plus  $Tv$ . And  $T$  acting on a times a vector is a times  $T$  of the vector. These two things make it into something we call a linear operator. It acts on the sum of vectors linearly and on a number times a vector. The number goes out and you act on the vector. So all you need to know for what a linear operator is, is how it acts on basis vectors. Because any vector on the vector space is a superposition of basis vectors.

So if you tell me how it acts on the basis vectors, you know everything. So we will figure out how the matrix representation of the operators arises from how it acts on the basis vectors. And you don't need an inner product.

The reason people think of this is they say, oh, the  $T_{ij}$  matrix element of  $T$  is the inner product of the operator between  $i$  and  $j$ . And this is true. But for that you need [? brass ?] and inner product, all these things. And they're not necessary. We'll define this without that. We don't need it.

So see you next time, and we'll continue that.

[APPLAUSE]

Thank you.