MITOCW | 18.06_L23

-- and lift-off on differential equations.

So, this section is about how to solve a system of first order, first derivative, constant coefficient linear equations.

And if we do it right, it turns directly into linear algebra.

The key idea is the solutions to constant coefficient linear equations are exponentials.

So if you look for an exponential, then all you have to find is what's up there in the exponent and what multiplies the exponential and that's the linear algebra.

So -- and the result -- one thing we will fine -- it's completely parallel to powers of a matrix.

So the last lecture was about how would you compute A to the K or A to the 100?

How do you compute high powers of a matrix?

Now it's not powers anymore, but it's exponentials.

That's the natural thing for differential equation.

Okay. But can I begin with an example?

And I'll just go through the mechanics.

How would I solve the differential -- two differential equations?

So I'm going to make it -- I'll have a two by two matrix and the coefficients are minus one two, one minus two and I'd better give you some initial condition.

So suppose it starts u at times zero -- this is u1, u2 -- let it -- let it -- suppose everything is in u1 at times zero.

So -- at -- at the start, it's all in u1.

But what happens as time goes on, du2/dt will -- will be positive, because of that u1 term, so flow will move into the u2 component and it will go out of the u1 component.

So we'll just follow that movement as time goes forward by looking at the eigenvalues and eigenvectors of that matrix.

That's a first job.

Before you do anything else, find the -- find the matrix and its eigenvalues and eigenvectors. So let me do that. Okay. So here's our matrix. Maybe you can tell me right away what -- what are the eigenvalues and -- eigenvalues anyway. And then we can check. But can you spot any of the eigenvalues of that matrix? We're looking for two eigenvalues. Do you see -- I mean, if I just wrote that matrix down, what -- what do you notice about it? It's singular, right. That -- that's a singular matrix. That tells me right away that one of the eigenvalues is lambda equals zero. I can -- that's a singular matrix, the second column is minus two times the first column, the determinant is zero, it's -- it's singular, so zero is an eigenvalue and the other eigenvalue will be -- from the trace. I look at the trace, the sum down the diagonal is minus three. That has to agree with the sum of the eigenvalue, so that second eigenvalue better be minus three. I could, of course -- I could compute -- why don't I over here -- compute the determinant of A minus lambda I, the determinant of this minus one minus lambda two one minus two minus lambda matrix. But we know what's coming.

When I do that multiplication, I get a lambda squared.

I get a two lambda and a one lambda, that's a three lambda.

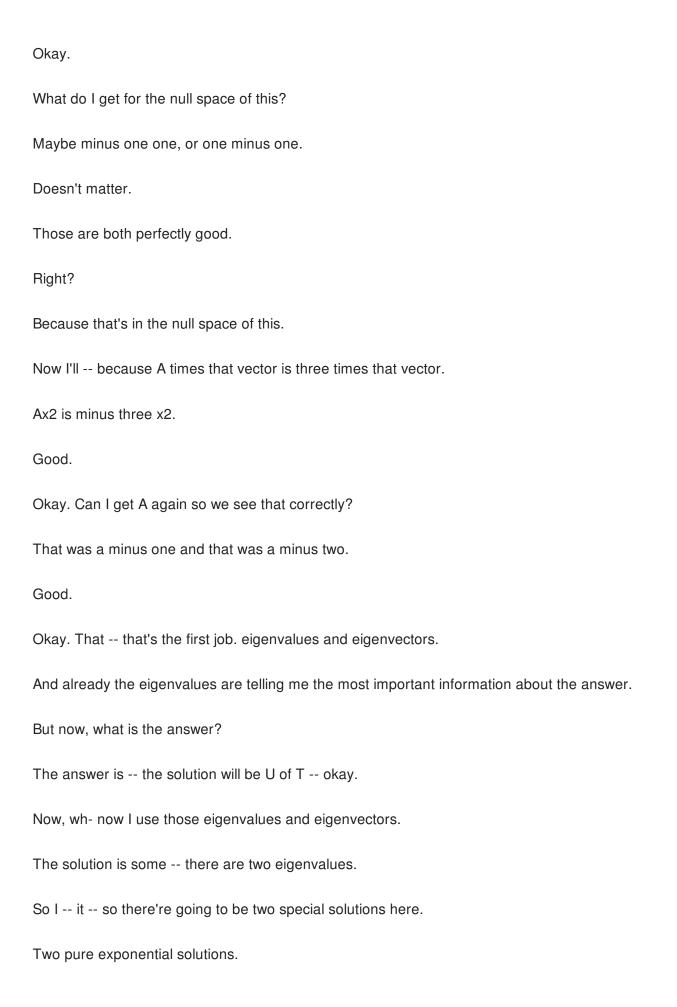
And then -- now I'm going to get the determinant, which is two minus two which is zero.

So there's my characteristic polynomial, this determinant.

And of course I factor that into lambda times lambda plus three and I get the two eigenvalues that we saw coming.

What else do I need? The eigenvectors. So before I even think about the differential equation or what -- how to solve it, let me find the eigenvectors for this matrix. Okay. So take lambda equals zero -- so that -- that's the first eigenvalue. Lambda one equals zero and the second eigenvalue will be lambda two equals minus three. By the way, I -- I already know something important about this. The eigenvalues are telling me something. You'll see how it comes out, but let me point to -- these numbers are -- this eigenvalue, a negative eigenvalue, is going to disappear. There's going to be an e to the minus three t in the answer. That e to the minus three t as times goes on is going to be very, very small. The other part of the answer will involve an e to the zero t. But e to the zero t is one and that's a constant. So I'm expecting that this solution'll have two parts, an e to the zero t part and an e to the minus three t part, and that -- and as time goes on, the second part'll disappear and the first part will be a steady It won't move, state. It will be -- at the end of -- as t approaches infinity, this part disappears and this is the -- the e to the zero t part is what I get. And I'm very interested in these steady states, so that's -- I get a steady state when I have a zero eigenvalue. Okay. What about those eigenvectors?

So what's the eigenvector that goes with eigenvalue zero?
Okay.
The matrix is singular as it is, the eigenvector is is the guy in the null space, so what vector is in the null space of that matrix?
Let's see.
I guess I probably give the free variable the value one and I realize that if I want to get zero I need a two up here
Okay? So Ax1 is zero x1.
A x1 is zero x1.
Fine. Okay.
What about the other eigenvalue?
Lambda two is minus three.
Okay.
How do I get the other eigenvalue, then?
For the moment can I mentally do it?
I subtract minus three along the diagonal, which means I add three can I I'll just do it with an erase erase for the moment.
So I'm going to add three to the diagonal.
So this minus one will become a two and I'll make it in big loopy letters and when I add three to this guy, the minus two becomes well, I can't make one very loopy, but how's that?
Okay. Now that's A minus three I A plus three I, sorry.
That's A plus three I.
It's supposed to be singular, right?
I if things if I did it right, this matrix should be singular and the x2, the eigenvector should be in its null space.



The first one is going to be either the lambda one tx1 and the -- so that solves the equation, and so does this one.

They both are solutions to the differential equation.

That's the general solution.

The general solution is a combination of that pure exponential solution and that pure exponential solution.

Can I just see that those guys do solve the equation?

So let me just check -- check on this one, for example.

I -- I want to check that the -- my equation -- let's

Check. remember, the equation -- du/dt is Au.

I plug in e to the lambda one t x1 and let's just see that the equation's okay.

I believe this is a solution to that equation.

So just plug it in.

On the left-hand side, I take the time derivative -- so the left-hand side will be lambda one, e to the lambda one t x1, right?

The time derivative -- this is the term that depends on time, it's just ordinary exponential, its derivative brings down a lambda one.

On the other side of the equation it's A times this

thing. A times either the lambda one t x one, and does that check out?

Do we have equality there?

Yes, because either the lambda one t appears on both sides and the other one is Ax1 equal lambda one x1 -- check.

Do you -- so, the -- we've come to the first point to remember.

These pure solutions.

Those pure solutions are the -- those pure exponentials are the differential equations analogue of -- last time we

had pure powers.

Last time -- so -- so last time, the analog was lambda -- lambda one to the K-th power x1, some amount of that, plus some amount of lambda two to the K-th power x2.

That was our formula from last time.

I put it up just to -- so your eye compares those two formulas.

Powers of lambda in the -- in the difference equation -- that -- this was in the -- this was for the equation uk plus one equals A uk.

That was for the finite step -- stepping by one.

And we got powers, now this is the one we're interested in, the exponentials.

So -- so that's -- that's the solution -- what are c1 and c2?

Then we're through.

What are c1 and c2?

Well, of course we know these actual things.

Let me just -- let me come back to this.

c1 is -- we haven't figured out yet, but e to the lambda one t, the lambda one is zero so that's just a one times x1 which is two one.

So it's c1 times this one that's not moving times the vector, the eigenvector two one and c2 times -- what's e to the lambda two t?

Lambda two is minus three.

So this is the term that has the minus three t and its eigenvector is this one minus one.

So this vector solves the equation and any multiple of it.

This vector solves the equation if it's got that factor e to the minus three t.

We've got the answer except for c1 and c2.

So -- so everything I've done is immediate as soon as you know the eigenvalues and eigenvectors. So how do we get c1 and c2? That has to come from the initial condition. So now I -- now I use -- u of zero is given as one zero. So this is the initial condition that will find c1 and c2. So let me do that on the board underneath. At t equals zero, then -- I get c1 times this guy plus c2 and now I'm at times zero. So that's a one and this is a one minus one and that's supposed to agree with u of zero one zero. Okay. That should be two equations. That should give me c1 and c2 and then I'm through. So what are c1 and c2? Let's see. I guess we could actually spot them by eye or we could solve two equations in two unknowns. Let's see. If these were both ones -- so I'm just adding -- then I would get three zero. So what's the -- what's the solution, then? If -- if c1 and c2 are both ones, I get three zero, so I want, like, one third of that, because I want to get one zero. So I think it's c1 equals a third, c2 equals a third. So finally I have the answer. Let me keep it in the -- in this board here. Finally the answer is one third of this plus one third of this.

Do you see what -- what's actually happening with this flow? This flow started out at -- the solution started out at one zero. Started at one zero. Then as time went on, people moved, essentially. Some fraction of this one moved here. And -- and in the limit, there's -- there's the limit, as -- right? As t goes to infinity, as t gets very large, this disappears and this is the steady state. So the steady state is -- so the steady state -- u -- we could call it u at infinity is one third of two and one. It's -- it's two thirds of one third. So that's the -- we really -- I mean, you're getting, like, total, insight into the behavior of the solution, what the differential equation does. Of course, we don't -- wouldn't always have a steady state. Sometimes we would approach zero. Sometimes we would blow up. Can we straighten out those cases? The eigenvalue should tell us. So when do we get -- so -- so let me ask first, when do we get stability? That's u of t going to zero. When would the solution go to zero no matter what the initial condition is? Negative eigenvalues, right. Negative eigenvalues. But now I have to -- I have to ask you for one more step.

Suppose the eigenvalues are complex numbers?

Because we know they could be.

Then we want stability -- this -- this -- we want -- we need all these e to the lambda t-s all going to zero and somehow that asks us to have lambda negative.

But suppose lambda is a complex number?

Then what's the test?

What -- if lambda's a complex number like, oh, suppose lambda is negative plus an imaginary part?

Say lambda is minus three plus six i?

What -- what happens then?

Can we just, like, do a -- a case here?

If -- if this lambda is minus three plus six it, how big is that number?

Does this -- does this imaginary part play a -- play a -- play a role here or not?

Or how big is -- what's the absolute value of that -- of that quantity?

It's just e to the minus three t, right?

Because this other part, this -- the -- the magnitude -- the -- this -- e to the six it -- what -- that has absolute value one.

Right?

That's just this cosine of six t plus i, sine of six t.

And the absolute value squared will be cos squared plus sine squared will be one.

This is -- this complex number is running around the unit

circle. This com- this -- the -- it's the real part that matters.

This is what I'm trying to do.

Real part of lambda has to be negative.

If lambda's a complex number, it's the real part, the minus three, that either makes us go to zero or doesn't, or let -- or blows up.

The imaginary part won't -- will just, like, oscillate between the two components.

Okay. So that's stability.

And what about -- what about a steady state?

When would we have, a steady state, always in the same direction?

So let me -- I'll take this part away -- when -- so that was, like, checking that it's -- that it's the real part that we care about.

Now, we have a steady state when -- when lambda one is zero and the other eigenvalues have what?

So I'm looking -- like, that example was, like, perfect for a steady state.

We have a zero eigenvalue and the other eigenvalues, we want those to disappear.

So the other eigenvalues have real part negative.

And we blow up, for sure -- we blow up if any real part of lambda is positive.

So if I -- if I reverse the sign of A -- of the matrix A -- suppose instead of the matrix I had, the A that I had, I changed it -- I changed all its sines.

What would that do to the eigenvalues and eigenvectors?

If I -- if -- if I know the eigenvalues and eigenvectors of A, tell me about minus A.

What happens to the eigenvalues?

For minus A, they'll all change sine.

So I'll have blow up.

This -- instead of the e to the minus three t, if I change that to minus -- if I change the sines in that matrix, I would change the trace to plus three, I would have an e to the plus three t and I would have blow up.

Of course the zero eigenvalue would stay at zero, but the other guy is taking off in -- if I reversed all the sines.

Okay. So this is -- this is the stability picture.

And for a two by two matrix, we can actually pin down even more closely what that means.

Can I -- let -- can I do that?

Let me do that -- I want to -- for a two by two matrix, I can tell whether the real part of the eigenvalues is negative, I -- well, let me -- let me tell you what I have in mind for that.

So two by two stability -- let me -- this is just a little comment.

Two by two stability.

So my matrix, now, is just a b c d and I'm looking for the real parts of both eigenvalues to be negative.

Okay.

What -- how can I tell from looking at the matrix, without computing its eigenvalues, whether the two eigenvalues are negative, or at least their real parts are negative?

What would that tell me about the trace?

So -- so the trace -- that's this a plus d -- what can you tell me about the trace in the case of a two by two stable matrix?

That means the eigenvalues have -- are negative, or at least the real parts of those eigenvalues are negative -- then, when I take the -- when I look at the matrix and find its trace, what -- what do I know about that?

It's negative, right.

This is the sum of the -- this equals -- this equals lambda one plus lambda two, so it's negative.

The two eigenvalues, by the way, will have -- if they're complex -- will have plus six i and minus six i.

The complex parts will -- will be conjugates of each other and disappear when we add and the trace will be negative.

Okay, the trace has to be negative.

Is that enough -- is a negative trace enough to make the matrix stable?

Shouldn't be enough, right? Can I -- can you make -- what's a matrix that has a negative trace but still it's not stable? So it -- it has a blow -- it still has a blow-up factor and a -- and a -- and a decaying one. So what would be a -- so just -- just to see -- maybe I just put that here. This -- now I'm looking for an example where the trace could be negative but still blow up. Of course -- yeah, let's just take one. Oh, look, let me -- let me make it minus two zero zero one. Okay. There's a case where that matrix has negative trace -- I know its eigenvalues of course. They're minus two and one and it blows up. It's got -- it's got a plus one eigenvalue here, so there would be an e to the plus t in the solution and it'll blow up if it has any second component at all. I need another condition. And it's a condition on the determinant. And what's that condition?

If I know that the two eigenvalues -- suppose I know they're negative, both negative.

What does that tell me about the determinant?

Let me ask again.

If I know both the eigenvalues are negative, then I know the trace is negative but the determinant is positive, because it's the product of the two eigenvalues.

So this determinant is lambda one times lambda two.

This is -- this is lambda one times lambda two and if they're both negative, the product is positive.

So positive determinant, negative trace.

I can easily track down the -- this condition also for the -- if -- if there's an imaginary part hanging around.

Okay. So that's a -- like a small but quite useful, because two by two is always important -- comment on stability.

Okay. So let's just look at the picture again.

Okay.

The main part of my lecture, the one thing you want to be able to, like, just do automatically is this step of solving the system.

Find the eigenvalues, find the eigenvectors, find the coefficients.

And notice -- what's the matrix -- in this linear system here, I can't help -- if I -- if I have equations like that -- that's my equations column at a time -- what's the matrix form of that equation?

So -- so this -- this system of equations is -- is some matrix multiplying c1, c2 to give u of zero.

One zero.

What's the matrix?

Well, it's obviously two one, one minus one, but we have a name, or at least a letter -- actually a name for that matrix.

Wh- what matrix are we s- are we -- are we dealing with here in this step of finding the c-s?

Its letter is S -- it's the eigenvector matrix.

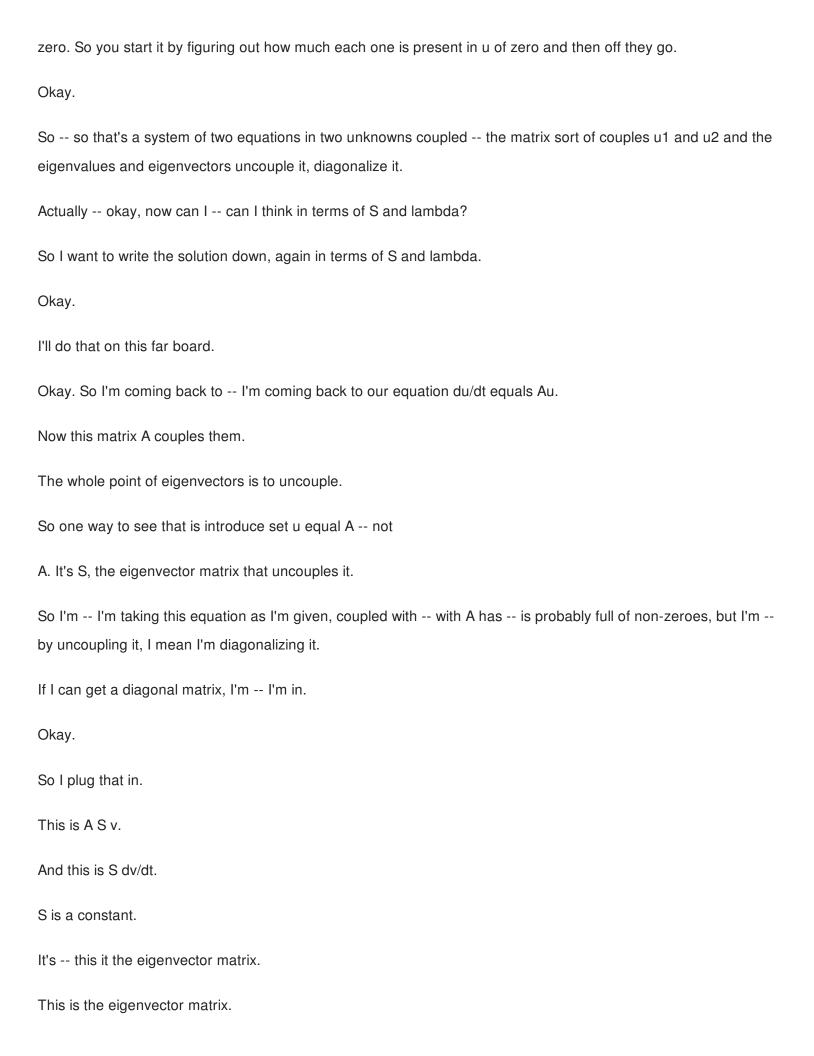
Of course.

These are the eigenvectors, there in the columns of our matrix.

So this is S c equals u of zero -- is the -- that step where you find the actual coefficients, you find out how much of each pure exponential is in the solution.

By getting it right at the start, then it stays right

forever. I -- you're seeing this picture that each -- each pure exponential goes on its own way once you start it from u of



Okay.

Now I'm going to bring S inverse over here.

And what have I got?

That combination S inverse A S is lambda, the diagonal matrix.

That's -- that's the point, that in -- if I'm using the eigenvectors as my basis, then my system of equations is just diagonal.

I -- each -- there's no coupling anymore -- dv1/dt is lambda one v1.

So let's just write that down. dv1/ dt is lambda one v1 and so on for all n of the equations.

It's a system of equations but they're not connected, so they're easy to solve and why don't I just write down the solution? v -- well, v is now some e to the lambda one t -- well, okay -- I guess my idea here now is to use, the natural notation -- v of T is e to the lambda tv of zero.

And u of t will be Se to the lambda t S inverse, u of zero.

This is the -- this is the, formula I'm headed for.

This -- this matrix, S e to the lambda t S inverse, that's my exponential.

That's my e to the At, is this Se to the lambdat Sinverse.

So my -- my job really now is to explain what's going on with this matrix up in the exponential.

What does that mean?

What does it mean to say e to a matrix?

This ought to be easier because this is e to a diagonal matrix, but still it's a matrix.

What do we mean by e to the A t?

Because really e to the A t is my answer here.

My -- my answer to this equation is -- this u of t, this is my -- this is my e to the At u of zero.

So it's -- my job is really now to say what's -- what does that mean?

What's the exponential of a matrix and why is that formula correct? Okay. So I'll put that on the board underneath. What's the exponential of a matrix? Let me come back here. So I'm talking about matrix exponentials. e to the At. Okay. How are we going to define the exponential of a -- of something? The trick -- the idea is -- the thing to go back to is exponential -- there's a power series for exponentials. That's how you -- you -- the good way to define e to the x is the power series one plus x plus one half x squared, one six x cubed and we'll do it now when the -- when we have a matrix. So the one becomes the identity, the x is At, the x squared is At squared and I divide by two. The cube, the x cube is At cubed over six, and what's the general term in here? At to the n-th power divided by -- and it goes on. But what do I divide by? So, you see the pattern -- here I divided by one, here I divided by one by two by six, those are the factorials. It's n factorial. That was, like, the one beautiful Taylor series.

That was, like, the one beautiful rayior series.

The one beautiful Taylor series -- well, there are two -- there are two beautiful Taylor series in this world.

The Taylor series for e to the x is the s with x to the n-th over n factorial.

And all I'm doing is doing the same thing for matrixes.

The other beautiful Taylor series is just the sum of x to the n-th not divided by n factorial.

Can you -- do you know what function that one is?

So if I take -- this is the series, all these sums are going from zero to infinity.

What's -- what function have I got -- this is like a side comment -- this is one plus x plus x squared plus x cubed plus x to the fourth not divided by anything, what's -- what's that function?

One plus x plus x squared plus x cubed plus x fourth forever is one over one minus x.

It's the geometric series, the nicest power series of all.

So, actually, of course, there would be a similar thing here.

If -- if I wanted, I minus A t inverse would be -- now I've got matrixes.

I've got matrixes everywhere, but it's just like that series with -- and just like this one without the divisions.

It's I plus At plus At squared plus At cubed and forever.

So that's actually a -- a reasonable way to find the inverse of a matrix.

If I chop it off -- well, it's reasonable if t is small.

If t is a small number, then -- then t squared is extremely small, t cubed is even smaller, so approximately that inverse is I plus At.

I can keep more terms if I like.

Do you see what I'm doing?

I'm just saying we can do the same thing for matrixes that we do for ordinary functions and the good thing about the exponential series -- so in a way, this series is better than this one.

Why?

Because this one always converges.

I'm dividing by these bigger and bigger numbers, so whatever matrix A and however large t is, that series -- these terms go to zero.

The series adds up to a finite sum, e to the At is a -- is -- is completely defined.

Whereas this second guy could fail, right?

If At is big -- somehow if At has an eigenvalue larger than one, then when I square it it'll have that eigenvalue squared, when I cube it the eigenvalue will be cubed -- that series will blow up unless the eigenvalues of At are smaller than one.

So when the eigenvalues of At are smaller than one -- so I'd better put that in.

The -- all eigenvalues of At below one -- then that series converges and gives me the inverse.

Okay. So this is the guy I'm chiefly interested in, and I wanted to connect it to -- oh, okay.

What's -- how do I -- how do I get -- this is my, like, main thing now to do -- how do I get e to the At -- how do I see that e to the At is the same as this?

In other words, I can find e to the At by finding S and lambda, because now e to the lambda t

is easy. Lambda's a diagonal matrix and we can write down either the lambda t -- and will right -- in a minute.

But how -- do you see what -- do you see that we're hoping for a -- we're hoping that we can compute either the A T from S and lambda -- and I have to look at that definition and say, okay, how do -- how do I get an S and the lambda to come out of that?

Okay, can -- do you see how I -- I want to connect that to that, from that definition.

So let me erase this -- the geometric series, which isn't part of the differential equations case and get the S and the lambda into this picture.

Oh, okay.

Here we go.

So identity is fine.

Now -- all right, you -- you -- you'll see how I'm -- how I going to get A replaced by S and S is in lambda's?

Well I use the fundamental formula of this whole chapter.

A is S lambda S inverse and then times t.





What's the exponential of that diagonal matrix?

Because our whole point is that this ought to be simple.

Our whole point is that to take the exponential of a diagonal matrix ought to be completely decoupled -- it ought to be diagonal, in other words, and it is.

It's just e to the lambda one t, e to the lambda two t, e to the lambda n t, all zeroes.

So -- so if we have a diagonal matrix and I plug it into this exponential formula, everything's diagonal and the diagonal terms are just the ordinary scaler exponentials e to the lambda one t.

Okay, so that's -- that's -- in a sense, I'm doing here, on this board, with -- with, like, formulas what I did on the -- in the first half of the lecture with specific matrix A and specific eigenvalues and eigenvectors.

The -- so let me show you the formulas again.

But the -- so you -- I guess -- what should you know from this

lecture? You should know what this matrix exponential is and, like, when does it go to zero?

Tell me again, now, the answer to that.

When does e to the At approach -- get smaller and smaller as t increases?

Well, the S and the S inverse aren't moving.

It's this one that has to get smaller and smaller and that one has this simple diagonal form.

And it goes to zero provided every one of these lambdas -- I -- I need to have each one of these guys go to zero, so I need every real part of every eigenvalue negative.

Right?

If the real part is negative, that's -- that takes the exponential -- that forces -- the exponential goes to zero.

Okay, so that -- that's really the difference.

If I can just draw the -- here's a picture of the -- of the -- this is the complex plain.

Here's the real axis and here's the imaginary axis.

And where do the eigenvalues have to be for stability in differential equations? They have to be over here, in the left half plain. So the left half plain is this plain, real part of lambda, less than zero. Where do the ma- where do the eigenvalues have to be for powers of the matrix to go to zero? Powers of the matrix go to zero if the eigenvalues are in here. So this is the stability region for powers -- this is the region -- absolute value of lambda, less than one. That's the stability for -- that tells us that the powers of A go to zero, this tells us that the exponential of A goes to zero. Okay. One final example. Let me just write down how to deal with a final example. Let's see. So my final example will be a single equation, u"+bu'+Ku=0. One equation, second order. How do I -- and maybe I should have used -- I'll use -- I prefer to use y here, because that's what you see in differential equations. And I want u to be a vector. So how do I change one second order equation into a two by two first order system? Just the way I did for Fibonacci. I'll let u be y prime and y. What I'm going to do is I'm going to add an extra equation, y prime equals y prime. So I take this -- so by -- so using this as the vector unknown, now my equation is u prime. My first order system is u prime, that'll be y double prime y prime, the derivative of u, okay, now the differential

