

Everything I say today is going to be for n -by- n systems, but for your calculations and the exams two-by-two will be good enough. Our system looks like that. Notice I am talking today about the homogeneous system, not the inhomogeneous system. So, homogeneous. And we have so far two basic methods of solving it. The first one, on which we spent the most time, is the method of where you calculate the eigenvalues of the matrix, the eigenvectors, and put them together to make the general solution. So eigenvalues, e -vectors and so on. The second method, which I gave you last time, I called "royal road," simply calculates the matrix e to the A and says that the solution is e to the A times x zero, the initial condition. That is very elegant. The only problem is that to calculate the matrix e to the A , although sometimes you can do it by its definition as an infinite series, most of the time the only way to calculate the matrix e to the A is by using the fundamental matrix. In other words, the normal way of doing it is you have to calculate it as the fundamental matrix time normalized at zero. So, as I explained at the end of last time and you practiced in the recitations, you have to find the fundamental matrix, which, of course, you have to do by eigenvalues and eigenvectors. And then you multiply it by its value at zero, inverse. And that, by magic, turns out to be the same as the exponential matrix. But, of course, there has been no gain in simplicity or no gain in ease of calculation. The only difference is that the language has been changed. Now, today is going to be devoted to yet another method which saves no work at all and only amounts to a change of language. The only reason I give it to you is because I have been begged by various engineering departments to do so -- -- because that is the language they use. In other words, each person who solves systems, some like to use fundamental matrices, some just calculate, some immediately convert the system by elimination into a single higher order equation because they are more comfortable with that. Some, especially if they are writing papers, they talk exponential matrices. But there are a certain number of engineers and scientists who talk decoupling, express the problem and the answer in terms of decoupling. And that is, therefore, what I have to explain to you today. So, the third method, today's method, I stress is really no more than a change of language. And I feel a little guilty about the whole business. Instead of going more deeply into studying these equations, what I am doing is like giving a language course and teaching you how to say hello and good-bye in French, German, Spanish, and Italian. It is not going very deeply into any of those languages, but you are going into the outside world, where people will speak these things. Here is an introduction to the language of decoupling in which for some people is the exclusive language in which they talk about systems. Now, I think the best way to, well, in a general way, what you try to do is as follows. You try to introduce new variables. You make a change of variables. I am going to do it two-by-two just to save a lot of writing out. And it's going to be a linear change of variables because we are interested in linear systems. The problem is to find u and v such that something wonderful happens, such that when you make the change of variables to express this system in terms of u and v it becomes decoupled. And that means the system turns into a system which looks like u prime equals k_1 times u and v prime equals k_2 times v . Such a system is called decoupled. Why? Well, a normal system is called coupled. Let's write out what it would be. Well, let's not write that. You know what it looks like. This is decoupled because it is not really a system at all. It is just two first-order equations sitting side by side and having nothing whatever to do with each other. This is two problems from the first day of the term. It is not one problem from the next to last day of the term, in other words. To solve this all you say is u is equal to some constant times e to the $k_1 t$ and v equals another constant times e to the $k_2 t$. Coupled means that the x and y occur in both equations on the right-hand side. And, therefore, you cannot solve separately for x and y , you must solve together for both of them. Here I can solve separately for u and v and, therefore, the system has been decoupled. Now, obviously, if you can do that it's an enormous advantage, not just to the ease of solution, because you can write down the solution immediately, but because something physical must be going on there. There must be some insight. There ought to be some physical reason for these new variables. Now, that is where I plan to start with. My plan for the lecture is first to work out, in some detail, a specific example where decoupling is done to show how that leads to the solution. And then we will go back and see how to do it in general -- -- because you will see, as I do the decoupling in this particular example, that that particular method, though it is suggested, will not work in general. I would need a more general method. But let's first go to the example. It is a slight modification of one you should have done in recitation. I don't think I worked one of these in the lecture, but to describe it I have to draw two views of it to make sure you know exactly what I am talking about. Sometimes it is called the two compartment ice cube tray problem, a very old-fashioned type of ice cube tray. Not a modern one that is all plastic where there is no leaking from one compartment to another. The old kind of ice cube trays, there were compartments and these were metal separated and you leveled the liquid because it could leak through the bottom that didn't go right to the bottom. If you don't know what I am talking about it makes no difference. This is the side view. This is meant to be twice as long. But, to make it quite clear, I will draw the top view of this thing. You have to imagine this is a rectangle, all the sides are parallel and everything. This is one and this is two. All I am trying to say is that the cross-sectional area of these two chambers, this one has twice the cross-sectional area of this one. So I will write a two here and I will write a one there. Of course, it is this hole here through which everything leaks. I am going to let x be the height of this liquid, the water here, and y the height of the water in that chamber. Obviously, as time goes by, they both reach the same height because of somebody's law. Now, what is the system of differential equations that controls this? Well, the essential thing is the flow rate through here. That flow rate through the hole in units, let's say, in liters per second. Just so you understand, I am talking about the volume of liquid. I am not talking about the velocity. That is proportional to the area of the hole. So the cross-sectional area of the hole. And it is also times the velocity of the flow, but the velocity of the flow depends upon the pressure difference. And that pressure difference depends upon the difference in height. All those are various people's laws. So times the height difference. Of course, you have to get the sign right. I have just pointed out the height difference is proportional to the pressure at the hole. And it is that pressure at the hole that determines the velocity with which the fluid flows through. Where does this all produce

our equations? Well, x' is equal to, therefore, some constant, depending on the area of the hole and this constant of proportionality with the pressure and the units and everything else times the pressure difference I am talking about. Well, if fluid is going to flow in this direction that must mean the y height is higher than the x height. So, to make x' positive, it should be $y - x$ here. Now, the y' is different. Because, again, the rate of fluid flow is determined. This time, if y' is positive, if this is rising, as it will be in this case, it's because the fluid is flowing in that direction. It is because x is higher than y . So this should be the same constant $x - y$. But notice that right-hand side is the rate at which fluid is flowing into this tank. That is not the rate at which y is changing. It is the rate at which $2y$ is changing. Why isn't there a constant here? There is. It's one. That is the one, this one cross-section. The area here is one and the cross-sectional area here is two. And that is the reason for the one here and the two here, because we are interested in the rate at which fluid is being added to this, which is only related to the height, the rate at which the height is rising if you take into account the cross-sectional area. So there is the system. In order to use nothing but integers here, I am going to take c equals to two, so I don't have to put in halves. The final system is $x' = x - 2x$, you have to write them in the correct order, and $y' = 2y - x$, the twos cancel because c is two, is $x - y$. So there is our system. Now the problem is I want to solve it by decoupling it. I want, in other words, to find new variables, u and v , which are more natural to the problem than the x and y that are so natural to the problem that the new system will just consistently be two side-by-side equations instead of the single equation. The question is, what should u and v be? Now, the difference between what I am going to do now and what I am going to do later in the period is later in the period I will give you a systematic way of finding what u and v should be. Now we are going to psyche out what they should be in the way in which people who solve systems often do. I am going to use the fact that this is not just an abstract system of equations. It comes from some physical problem. And I ask, is there some system of variables, which somehow go more deeply into the structure of what's going on here than the naive variables, which simply tell me how high the two tank levels are? That is the obvious thing I can see, but there are some variables that go more deeply. Now, one of them is sort of obvious and suggested both the form of the equation and by this. Simply, the difference in heights is, in some ways, a more natural variable because that is directly related to the pressure difference, which is directly related to the velocity of flow. They will differ by just constant. I am going to call that the second variable, or the difference in height let's call it. That's $x - y$. That is a very natural variable for the problem. The question is, what should the other one be? Now you sort of stare at that for a while until it occurs to you that something is constant. What is constant in this problem? Well, the tank is sitting there, that is constant. But what thing, which might be a variable, clearly must be a constant? It will be the total amount of water in the two tanks. These things vary, but the total amount of water stays the same because it is a homogenous problem. No water is coming in from the outside, and none is leaving the tanks through a little hole. Okay. What is the expression for the total amount of water in the tanks? $x + 2y$. Therefore, that is a natural variable also. It is independent of this one. It is not a simple multiply of it. It is a really different variable. This variable represents the total amount of liquid in the two tanks. This represents the pressure up to a constant factor. It is proportional to the pressure at the hole. Okay. Now what I am going to do is say this is my change of variable. Now let's plug in and see what happens to the system when I plug in these two variables. And how do I do that? Well, I want to substitute and get the new system. The new system, or rather the old system, but what makes it new is in terms of u and v . What will that be? Well, $u' = x' + 2y'$. But I know what $x' + 2y'$ is because I can calculate it for this. What will it be? $x' + 2y' = x - 2x + 2(2y - x) = x - 2x + 4y - 2x = -3x + 4y$, which is zero. And how about these two? $2y' - x'$, because I want this plus twice that, so it's $2y' - x'$, again, zero. The right-hand side becomes zero after I calculate $x' + 2y'$. So that is zero. That would just, of course, clear. Now, that makes sense, of course. Since the total amount is constant, that says that u' is zero. Okay. What is v' ? $v' = x' - y'$. What is that? Well, once again we have to calculate. $x' - y' = x - 2x - (2y - x) = x - 2x - 2y + x = -2y$, which is $-3x$, and $2y' - x' = 4y - 2x - (x - 2x) = 4y - 2x - x + 2x = 4y - x$, which makes $3y$. All right. What is the system? The system is $u' = 0$ and $v' = -3x$ and $3y$. But $x - y$ is v . In other words, these new two variables decouple the system. And we got them, as scientists often do, by physical considerations. These variables go more deeply into what is going on in that system of two tanks than simply the two heights, which are too obvious as variables. All right. What is the solution? Well, the solution is, u equals a constant and v is equal to? Well, the solution to this equation is a different arbitrary constant from that one. These are side-by-side equations that have nothing whatever to do with each other, remember? $u = c_1 e^{-3t}$. Now, there are two options. Either one leaves the solution in terms of those new variables, saying they are more natural to the problem, but sometimes, of course, one wants the answer in terms of the old one. But, if you do that, then you have to solve that. In order to save a little time, since this is purely linear algebra, I am going to write -- Instead of taking two minutes to actually do the calculation in front of you, I will just write down what the answer is -- in terms of u and x and y . In other words, this is a perfectly good way to leave the answer if you are allowed to do it. But if somebody says they want the answer in terms of x and y , well, you have to give them what they are paying for. In terms of x and y , you have first to solve those equations backwards for x and y in terms of u and v in which case you will get $x = \frac{1}{3}u + 2v$. Use the inverse matrix or just do elimination, whatever you usually like to do. And the other one will be one-third of u minus v . And then, if you substitute in, you will see what you will get is one-third of c_1 . Sorry. u is c_1 . $c_1 + 2c_2 e^{-3t}$. And this is one-third of $c_1 - c_2 e^{-3t}$. And so, the final solution is, in terms of the way we usually write out the answer, x will be what? Well, it will be one-third c_1 times the eigenvector one, one plus one-third times c_2 times the eigenvector two, negative one times e^{-3t} . That is the solution written out in terms of x and y either as a vector in the usual way or separately in terms of x and y . But, notice, in order to do that you have to have these backwards equations. In other words, I need the equations in that form. I need the equations because they tell me what the new variables are. But I also have to have the equations the other way in order to get the solution in terms of x and v . finally. Okay. That was all an example. For the rest of the period. I

would like to show you the general method of doing the same thing which does not depend upon being clever about the choice of the new variables. And then, at the very end of the period, I will apply the general method to this problem to see whether we get the same answer or not. What is the general method? Our problem is the decouple. Now, the first thing is you cannot always decouple. To decouple the eigenvalues must all be real and non-defective. In other words, if they are repeated they must be complete. You must have enough independent eigenvectors. So they must be real and complete. If repeated, they must be complete. They must not be defective. As I told you at the time when we studied complete and incomplete, the most common case in which this occurs is when the matrix is symmetric. If the matrix is real and symmetric then you can always decouple the system. That is a very important theorem, particularly since many of the equilibrium problems normally lead to symmetric matrices and are solved by decoupling. Okay. So what are we looking for? We are assuming this and we need it. In general, otherwise, you cannot decouple if you have complex eigenvalues and you cannot decouple if you have defective eigenvalues. Well, what are we looking for? We are looking for new variables. u, v equals a_1, b_1, a_2, b_2 times the x, y . And this matrix is called D , the decoupling matrix and is what we are looking for. How do I choose those new variables u and v when I don't have any physical considerations to guide me as I did before? Now, the key is to look instead at what you are going to need. Remember, we are changing variables. And, as I told you from the first days of the term, when you change variables look at what you are going to need to substitute in to make the change of variables. Don't just start writing equations. What we are going to need to plug into that system and change it to the (u, v) coordinates is not u and v in terms of x and y . What we need is x and y in terms of u and v to do the substitution. What we need is the inverse of this. So, in order to do the substitution, what we need is (x, y) . Oops. Let's call them prime. Let's call these a_1, b_1, a_2, b_2 because these are going to be much more important to the problem than the other ones. Okay. I am going to, I should call this matrix D inverse, that would be a sensible thing to call it. Since this is the important matrix, this is the one we are going to need to do the substitution, I am going to give it another letter instead. And the letter that comes after D is E . Now, E is an excellent choice because it is also the first letter of the word eigenvector. And the point is the matrix E , which is going to work, is the matrix whose columns are the two eigenvectors. The columns are the two eigenvectors. Now, even if you didn't know anything that would be practically the only reasonable choice anybody could make. What are we looking for? To make a linear change of variables like this really means to pick new i and j vectors. You know, from the first days of 18.02, what you want is a new coordinate system in the plane. And the coordinate system in the plane is determined as soon as you tell what the new i is and what the new j is in the new system. To establish a linear change of coordinates amounts to picking two new vectors that are going to play the role of i and j instead of the old i and the old j . Okay, so pick two vectors which somehow are important to this matrix. Well, there are only two, the eigenvectors. What else could they possibly be? Now, what is the relation? I say with this, what happens is I say that α_1 corresponds, and α_2 , these are vectors in the xy -system. Well, if I change the coordinates to u and v , in the uv -system they will correspond to the vectors one, zero. In other words, the vector that we would normally call i in the u, v system. And this one will correspond to the vector zero, one. Now, if you don't believe that I will calculate it for you. The calculation is trivial. Look. What have we got? (x, y) equals a_1, b_1, a_2, b_2 . This is the column vector α_1 . This is the column vector α_2 . Now, here is u and v . Suppose I make u and v equal to one, zero, what happens to x and y ? Your matrix multiply. One, zero. So a_1 plus zero, b_1 plus zero. It corresponds to the column vector (a_1, b_1) . And in the same way zero, one corresponds to (a_2, b_2) . Just by matrix multiplication. And that shows that these correspond. In the uv -system the two eigenvectors are now called i and j . Well, that looks very promising, but the program now is to do the substitution to substitute into the system x prime equals Ax and see if it is decoupled in the uv -coordinates. Now, I don't dare let you do this by yourself because you will run into trouble. Nothing is going to happen. You will just get a mess and will say I must be missing something. And that is because you are missing something. What you are missing, and this is a good occasion to tell you, is that, in general, three-quarters of the civilized world does not introduce eigenvalues and eigenvectors the way you learn them in 18.03. They use a different definition that is identical. I mean it is equivalent. The concept is the same, but it looks a little different. Our definition is what? Well, what is an eigenvalue and eigenvector? The basic thing is this equation. This is a two-by-two matrix, right? This is a column vector with two entries. The product has to be a column vector with two entries, but both entries are supposed to be zero so I will write it this way. This way first defines what an eigenvalue is. It is something that makes the determinant zero. And then it defines what an eigenvector is. It is, then, a solution to the system that you can get because the determinant is zero. Now, that is not what most people do. What most people do is the following. They write this equation differently by having something on both sides. Using the distributive law, what goes on the left side is $A \alpha_1$. What is that? That is a column vector with two entries. What goes on the right? Well, λ_1 times the identity times α_1 . Now, the identity matrix times anything just reproduces what was there. There is no difference between writing the identity times α_1 and just α_1 all by itself. So that is what I am going to do. This is the definition of eigenvalue and eigenvector that all the other people use. Most linear algebra books use this definition, or most books use a different approach and say, here is an eigenvalue and an eigenvector. And it requires them to define them in the opposite order. First what α_1 is and then what λ_1 is. See, I don't have any determinant now. So what is the definition? And they like it because it has a certain geometric flavor that this one lacks entirely. This is good for solving differential equations, which is why we are using it in 18.03, but this has a certain geometric content. This way thinks of A as a linear transformation of the plane, a shearing of the plane. You take the plane and do something to it. Or, you squish it like that. Or, you rotate it. That's okay, too. And the matrix defines a linear transformation to the plane, every vector goes to another vector. The question it asks is, is there a vector which is taken by this linear transformation and just left alone or stretched, is kept in the same direction but stretched? Or, maybe its direction is reversed and it is stretched or it shrunk. But, in general, if there are real eigenvalues there will be such vectors that are just left in the same direction but just stretched or shrunk. And what is the λ ? The λ then is the amount by

which they are stretched or shrunk, the factor. This way, first we have to find the vector, which is left essentially unchanged, and then the number here that goes with it is the stretching factor or the shrinking factor. But the end result is the pair α_1 and λ_1 , regardless of which order you find them, satisfied the same equation. Now, a consequence of this definition we are going to need in the calculation that I am going to do in just a moment. Let me calculate that out. What I want to do is calculate the matrix A times E . I am going to need to calculate that. Now, what is that? Remember, E is the matrix whose columns are the eigenvectors. That is the matrix α_1, α_2 . Now, what is this? Well, in both Friday's lecture and Monday's lecture, I used the fact that if you do a multiplication like that it is the same thing as doing the multiplication $A \alpha_1$ and putting it in the first column. And then $A \alpha_2$ is the column vector that goes in the second column. But what is this? This is $\lambda_1 \alpha_1$. And this is $\lambda_2 \alpha_2$ by this other definition of eigenvalue and eigenvector. And what is this? Can I write this in terms of matrices? Yes indeed I can. This is the matrix α_1, α_2 times this matrix $\lambda_1, \lambda_2, 0, 0$. Check it out. $\lambda_1 \alpha_1 + 0, \lambda_2 \alpha_2 + 0$ this thing plus zero, the first entry is exactly that. And the same way the second column doing the same calculation is exactly this. What is that? That is e times this matrix $\lambda_1, 0, 0, \lambda_2$. Okay. We are almost finished now. Now we can carry out our work. We are going to do the substitution. I start with a system. Remember where we are. I am starting with this system. I am going to make the substitution x equal to this matrix E , whose columns are the eigenvectors. I am introducing, in other words, new variables u and v according to that thing. u is the column vector, u and v . And x , as usual, is the column vector x and y . So I am going to plug it in. Okay. Let's plug it in. What do I get? I take the derivative. E is a constant matrix so that makes E times u' , is equal to A times x is E times u again. Now, at this point, you would be stuck, except I calculated for you A times E is E times that funny diagonal matrix with the lambdas. So this is E times that funny matrix of the lambdas, the eigenvalues, and still the u at the end of it. So where are we? E times u' equals E times this thing. Well, multiply both sides by E^{-1} and you can cancel them out. And so the end result is that after you have made the substitution in terms of the new variables u , what you get is $u' = \lambda_1 u, \lambda_2 v, 0, 0$. Let's write that out in terms of a system. This is $u' = \lambda_1 u, v'$ here. It is $\lambda_1 u + 0 v$. And the other one is $v' = 0 u + \lambda_2 v$. We are decoupled. In just one sentence you would say -- In other words, if you were reading a book that sort of assumed you knew what was going on, all it would say is as usual. That is to make you feel bad. Or, as is well-known to make you feel even worse. Or, the system is decoupled by choosing as the new basis for the system the eigenvectors of the matrix and in terms of the resulting new coordinates, the decoupled system will be the following where the constants are the eigenvalues. And so the solution will be $u = c_1 e^{\lambda_1 t}$ and $v = c_2 e^{\lambda_2 t}$. Of course, if you want it back in terms now of x and y , you will have to go back to here, to these equations and then plug in for u and v what they are. And then you will get the answer in terms of x and y . Okay. We have just enough time to actually carry out this little program. It takes a lot longer to derive than it does actually to do, so let's do it for this system that we were talking about before. Decouple the system, $x', y' = \begin{pmatrix} -2 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$. Okay. What do I do? Well, I first have to calculate the eigenvalues in the eigenvectors, so the $E \lambda$'s and $E v$'s. The characteristic equation is $\lambda^2 - \text{trace}(\lambda) + \det = 0$. The trace is negative three, but you have to change the sign. The determinant is two minus two, so that is zero. There is no constant term here. It is zero. That is the characteristic equation. The roots are obviously $\lambda_1 = 0, \lambda_2 = -3$. And what are the eigenvectors that go with that? With $\lambda_1 = 0$ goes the eigenvector, minus two. Well, I subtract zero here, so the equation I have to solve is $-2a_1 + 2a_2 = 0$. I am not going to write $2a_2$, which is what you have been writing up until now. The reason is because I ran into trouble with the notation and I had to use, as the eigenvector, not a_1, a_2 but a_1, b_1 . So it should be $a_1 = b_1$ here, not the a_2 that you are used to. The solution, therefore, is $\alpha_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. And for $\lambda_2 = -3$, the corresponding eigenvector this time will be -- Now I have to subtract negative three from here, so negative two minus negative three makes one. That is $a_1 + 2b_1 = 0$, a logical choice for the eigenvector here. The second eigenvector would be make $b_1 = 1$, let's say, and then $a_1 = -2$. Okay. Now what do we have to do? Now, what we want is the matrix E . The matrix E is the matrix of eigenvectors, so it is the matrix $\begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}$. The next thing we want is what the new variables u and v are. For that, we will need E^{-1} . How do you calculate the inverse of a two-by-two matrix? You switch the two diagonal elements, there I have switched them, and you leave the other two where they are but change their sign. So it is two up here and negative one there. Maybe I should make this one purple and then that one purple to indicate that I have switched them. I am not done yet. I have to divide by the determinant. What is the determinant? It is one minus negative two, which is three, so I have to divide by three. I multiply everything here by one-third. Okay. And what is the decoupled system? The new variables are $u = \frac{1}{3}(x + 2y), v = \frac{1}{3}(-x + y)$. In other words, the new variables are given by D . It is $u, v = \frac{1}{3} \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$. That is the expression for u, v in terms of x and y . It's this matrix D , the decoupling matrix which is the one that is used. And that gives this system $u' = \lambda_1 u, v' = \lambda_2 v$. And what is the v entry? v is one-third of $-x + y$. Now, are those the same variables that I used before? Yes. This is my new and better you, the one I got by just blindly following the method instead of looking for physical things with physical meaning. It differs from the old one just by a constant factor. Now, that doesn't have any effect on the resulting equation because if the old one is $u' = \lambda_1 u$ the new one is $\frac{1}{3}u' = \lambda_1 \frac{1}{3}u$. It is still the same equation, in other words. And how about this one? This one differs from the other one by the factor minus one-third. If I multiply that v through by minus one-third, I get this v . And, therefore, that too does not affect the second equation. I simply multiply both sides by minus one-third. The new v still satisfies the equation $v' = \lambda_2 v$.