PROFESSOR STRANG:

Starting with a differential equation. So key point here in this lecture is how do you start with a differential equation and end up with a discrete problem that you can solve? But simple differential equation. It's got a second derivative and I put a minus sign for a reason that you will see. Second derivatives are essentially negative definite things so that minus sign is to really make it positive definite. And notice we have boundary conditions. At one end the solution is zero, at the other end it's zero. So this is fixed-fixed. And it's a boundary value problem. That's different from an initial value problem. We have x space, not time. So we're not starting from some thing and oscillating or growing or decaying in time. We have a fixed thing. Think of an elastic bar. An elastic bar fixed at both ends, maybe hanging by its own weight. So that load f(x) could represent the weight.

Maybe the good place to sit is over there, there are tables just, how about that? It's more comfortable.

So we can solve that equation. Especially when I change f(x) to be one. As I plan to do. So I'm going to change, I'm going to make it-- it's a uniform bar because there's no variable coefficient in there and let me make it a uniform load, just one. So it actually shows you that, I mentioned differential equations and we'll certainly get onto Laplace's equation, but essentially our differential equations will not-- this isn't a course in how to solve ODEs or PDEs. Especially not ODEs. It's a course in how to compute solutions. So the key idea will be to replace the differential equation by a difference equation. So there's the difference equation. And I have to talk about that. That's the sort of key point, that-- up here you see what I would call a second difference. Actually with a minus sign. And on the right-hand side you see the load, still f(x).

Can I move to this board to explain differences? Because this is like, key step is given the differential equation replace it by difference equation. And the interesting point is you have many choices. There's one differential equation but even for a first derivative, so this, if you remember from calculus, how did you start with the derivative? You started by something before going to the limit. h or delta x goes to zero in the end to get the derivative. But this was

a finite difference. You moved a finite amount. And this is the one you always see in calculus courses. $u(x+h) - u(x)$, just how much did that step go. You divide by the delta x, the h, and that's approximately the derivative, $u'(x)$. Let me just continue with these others. I don't remember if calculus mentions a backward difference. But you won't be surprised that another possibility, equally good more or less, would be to take the point and the point before, take the difference, divide by delta x. So again all these approximate u'.

And now here's one that actually is really important. A center difference. It's the average of the forward and back. If I take that plus that, the $u(x)$'s cancel and I'm left with, I'm centering it. This idea of centering is a good thing actually. And of course I have to divide by 2h because this step is now 2h, two delta x's. So that again is going to represent u'. But so we have a choice if we have a first derivative. And actually that's a big issue. You know, one might be called upwind, one might be called downwind, one may be called centered. It comes up constantly in aero, in mechanical engineering, everywhere. You have these choices to make. Especially for the first difference. We don't have, I didn't allow a first derivative in that equation because I wanted to keep it symmetric and first derivatives, first differences tend to be anti-symmetric. So if we want to get our good matrix K, and I better remember to divide by h squared, because the K just has those that I introduced last time and will repeat, the K just has the numbers -1, 2, -1.

Now, first point before we leave these guys what's up with them? How do we decide which one is better? There's something called the order of accuracy. How close is the difference to the derivative? And the answer is the error is of size h. So I would call that first order accurate. And I can repeat here but the text does it, how you recognize what-- This is the sort of local error, truncation error, whatever, you've chopped off the exact answer and just did differences. This one is also order h. And in fact, the h terms, the leading error, which is going to multiply the h, has opposite sign in these two and that's the reason centered differences are great. Because when you average them your center things-- this is correct to order h squared. And I may come back and find out why that h squared term is.

Maybe I'll do that. Yeah. Why don't I just? Second differences are so important. Why don't we just see. And centered differences. So let me see. How do you figure $u(x+h)$? This is a chance to remember something called Taylor series. But that was in calculus. If you forgot it, you're a normal person. So but what does it say? That's the whole point of calculus, in a way. That if I move a little bit, I start from the point x and then there's a little correction and that's given by

the derivative and then there's a further correction if I want to go further and that's given by half of h squared, you see the second order correction, times the second derivative. And then, of course, more. But that's all you ever have to remember. It's pretty rare. Second order accuracy is often the goal in scientific computing. First order accuracy is, like, the lowest level. You start there, you write a code, you test it and so on. But if you want production, if you want accuracy, get to second order if possible.

Now, what about this u(x-h)? Well, that's a step backwards, so that's u(x), now the step is -h, but then when I square that step I'm back to +h^2 u''(x) and so on. Ooh! Am I going to find even more accuracy? I could tell you what the next term is. Plus h^3 upon 6, that's 3*2*1, u triple-prime. And this would be, since this step is -h now, it would be -h^3 / 6 u'''. So what happens when I take the difference of these two? Remember now that centered differences subtract this from this. So now that u(x+h) - u(x-h) is zero, 2hu' subtracting that from that is zero, two of these, so I guess that we really have an h^3 / 3 u'''. And now when I divide by the 2h, can I just divide by 2h here? Oh yeah, it's coming out right, divide by 2h, divide this by 2h, that'll make it an h squared over six.

I've done what looks like a messy computation. I'm a little sad to start a good lecture, important lecture by such grungy stuff. But it makes the key point. That the centered difference gives the correct derivative with an error of order h^2. Where the error for the first differences, the h would have been there. And we can test it. Actually we'll test it. Okay for that? This is first differences. And that's a big question: what do you replace the first derivative by if there is one? And you've got these three choices. And usually this is the best choice.

Now to second derivatives. Because our equation has got u'' in it. So what's a second derivative? It's the derivative of the derivative. So what's the second difference? It's the difference, first difference of the first difference. So the second difference, the natural second difference would be-- so now let me use this space for second differences. Second differences. I could take the forward difference of the backward difference. Or I could take the backward difference of the forward difference. Or you may say why don't I take the centered difference of the centered difference. All those, in some sense it's delta squared, but which to take? Well actually those are the same and that's the good choice, that's the 1, -2, 1 choice. So let me show you that. Let me say what's the matter with that. Because now having said how great centered differences are, first differences, why don't I just repeat them for second differences?

Well the trouble is, let me say in a word without even writing, well I could even write a little, the center difference, suppose I'm at a typical mesh point here. The center difference is going to take that value minus that value. But then if I take the center difference of that I'm going to be out here. I'm going to take this value, this value, and this value. I'll get something correct. Its accuracy will be second order, good. But it stretches too far. We want compact difference molecules. We don't want this one, minus two of this, plus one of that. So this would give us a 1, 0, -2 , 0, 1. I'm just saying this and then I'll never come back to it because I don't like this one, these guys give 1, -2, 1 without any gaps. And that's the right choice. And that's the choice made here.

So I'm not thinking you can see it in your head, that the difference of the difference-- But well, you almost can. If I take this, yeah. Can you sort of see this without my writing it? If I take the forward difference and then I subtract the forward difference to the left, do you see that I'll have minus two. So there is what I started with. I subtract u(x) - U(x-h) and I get two -u(x)'s This is what I get. Now I'm calling that u_i. I better make completely clear about the minus sign. The forward difference or the backward difference, what this leads is 1, -2, 1. That's the second difference. Very important to remember, the second difference of a function is the function, the value ahead, minus two of the center, plus one of the left. It's centered obviously, symmetric, right? Second differences are symmetric. And because I want a minus sign I want minus the second difference and that's why you see here -1, 2, -1. Because I wanted positive two there. Are you okay? This is the natural replacement for -u''. And I claim that this second difference is like the second derivative, of course.

And why don't we just check some examples to see how like the second derivative it is. So I'm going to take the second difference of some easy functions. It's very important that these come out so well. So I'm going to take the second difference. I'm going to write it as sort of a matrix. So this is like the second difference. Yeah, because this is good. I'm inside the region, here. I'm not worried about the boundaries now. Let me just think of myself as inside. So I have second differences and suppose I'm applying it to a vector of all ones. What answer should I get? So if I think of calculus, it's the second derivative of one, of the constant function. So what answer am I going to get? Zero. And do I get zero? Of course. I get zero. Right? All these second differences are zero. Because I'm not worrying about the boundary yet. So that's like, check one. It passed that simple test.

Now let me move up from constant to linear. And so on. So let me apply second differences to

a vector that's growing linearly. What answer do I expect to get for that? So remember I'm doing second differences, like second derivatives, or minus second derivatives, actually. So what do second derivatives do to a linear function? If I take a straight line I take the-- sorry, second derivatives. If I take second derivatives of a linear function I get? Zero, right. So I would hope to get zero again here and I do. Right? -1+4-3=0. Minus one, sorry, let me do it here, -2+6-4. And actually, that's consistent with our little Taylor series stuff. The function x should come out right.

Now what about-- now comes the moment. What about x squared? So I'm going to put squares in now. Do I expect to get zeroes? I don't think so. Because let me again test it by thinking about what second derivative. So now I'm sort of copying second derivative of x squared, which is? Second derivative of x squared is? Two, right? First derivative's 2x, second derivative is just two. So it's a constant. And remember I put in a minus sign so I'm wondering, do I get the answer minus two? All the way down. -4+8-9. Whoops. What's that? What do I get there? What do I get from that second difference of these squares? -4+8-9 is? Minus two, good. So can we keep going? -4+18-16. What's that? -4+18-16, so I've got -20+18. I got minus two again. -9, 32, -25, it's right. The second differences of the vector of squares, you could say, is a constant vector with the right number. And that's because that second difference is second order accurate. It not only got constants right and linears right, it got quadratics right. So that's, you're seeing second differences. We'll soon see that second differences are also on the ball when you apply them to other vectors. Like vectors of sines or vectors of cosines or exponentials, they do well. So that's just a useful check which will help us over here.

Okay, can I come back to the part of the lecture now? Having prepared the way for this. Well, let's start right off by solving the differential equation. So I'm bringing you back years and years and years, right? Solve that differential equation with these two boundary conditions. How would you do that in a systematic way? You could almost guess after a while, but systematically if I have a linear, I notice-- What do I notice about this thing? It's linear. So what am I expecting? I'm expecting, like, a particular solution that gives the correct answer, one, and some null space solution or whatever I want to call it, homogeneous solution that gives zero and has some arbitrary constants in it. Give me a particular solution. So this is going to be our answer. This'll be the general solution to this differential equation. What functions have minus the second derivative equal one, that's all I'm asking. What are they? So what is one of them? One function that has its second derivative is a constant and that constant is minus

one. So if I want the second derivative to be a constant, what am I looking at? x squared. I'm looking at x squared. And I just want to figure out how many x squareds to get a one. So some number of x squareds and how many do I want? -1/2, good. Good. -1/2. Because x squared would give me two but I want minus one so I need -1/2. Okay that's the particular solution.

Now throw in all the solutions, I can add in any solution that has a zero on the right side, so what functions have second derivatives equals zero? x is good. I'm looking for two because it's a second derivative, second order equation. What's the other guy? Constant, good. So let me put the constant first, C, say, and Dx. Two constants that I can play with and what use am I going to make of them? I'm going to use those to satisfy the two boundary conditions. And it won't be difficult. You could say plug in the first boundary condition, get an equation for the constants, plug in the second, got another equation, we'll have two boundary conditions, two equations, two constants. Everything's going to come out. So if I plug in u(0)=0, what do I learn? C is zero, right? If I plug in, is that right? If I plug in zero, then that's zero already, this is zero already, so I just learned that C is zero. So C is zero. So I'm down to one constant, one unused boundary condition. Plug that in. u(1) = -1/2. What's D? It's 1/2, right. D is 1/2. So can I close this up? There's 1/2. Dx is 1/2. Now it just always pays to look back. At x=0, that's obviously zero. At x=1 it's zero because those are the same and I get zero. So -1/2 x squared plus 1/2 x. That's the kind of differential equation and solution that we're looking for. Not complicated nonlinear stuff.

So now I'm ready to move to the difference equation. So again, this is a major step. I'll draw a picture of this from zero to one. And if I graph that I think I get a parabola, right? A parabola that has to go through here. So it's some parabola like that. That would be always good, to draw a graph of the solution. Now, what do I get here? Moving to the difference equation. So that's the equation, and notice its boundary conditions. Those boundary conditions just copied this one because I've chopped this up. I've got i = 1, 2, 3, 4, 5, and this is one, the last point then is 6h. h is 1/6. What's going to be the size of my matrix and my vector and my unknown u here? How many unknowns am I going to have? Let's just get the overall picture right. What are the unknowns going to be? They're going to be u_1, u_2, u_3, u_4, u_5. Those are unknown. Those will be some values, I don't know where, maybe something like this because they'll be sort of like that one. And this is not an unknown, u_6; this is not an unknown, u_0; those are the ones we know. So this is what the solution to a difference equation looks like. It gives you a discrete set of unknowns. And then, of course MATLAB or any code could connect them up by straight lines and give you a function. But the heart of it is these five values. Okay,

good. And those five values come from these equations. I'm introducing this subscript stuff but I won't need it all the time because you'll see the picture. This equation applies for i equal one up to five. Five inside points and then you notice how when i is one, this needs u_0, but we know u_0. And when i is five, this needs u_6, but we know u_6. So it's a closed five by five system and it will be our matrix. That -1, 2, -1 is what sits on the matrix. When we close it with the two boundary conditions it chops off the zero column, you could say and chops off the six column and leaves us with a five by five problem and yeah.

I guess this is a step not to jump past because it takes a little practice. You see I've written the same thing two ways. Let me write it a third way. Let me write it out clearly. So now here I'm going to complete this matrix with a two and a minus one and a two and a minus one and now it's five by five. And those might be u, but I don't know if they are, so let me put in u_1, u_2, u_3, u_4, and u_5. Oh and divide by h squared. I'll often forget that. So I'm asking you to see something that if you haven't, after you get the hang of it it's like, automatic. But I have to remember it's not automatic. Things aren't automatic until you've done them a couple of times. So do you see that that is a concrete statement of this? This delta x squared is the h squared. And do you see those differences when I do that multiplication that they produce those differences? And now, what's my right-hand side? Well I've changed the right-hand side to one to make it easy. So this right-hand side is all ones. And this is the problem that MATLAB would solve or whatever code. Find a difference code. I've got to a linear system, five by five, it's fortunately-- the matrix is not singular, there is a solution.

How does MATLAB find it? It does not find it by finding the inverse of that matrix. Monday's lecture will quickly review how to solve five equations and five unknowns. It's by elimination, I'll tell you the key word. And that's what every code does. And sometimes you would have to exchange rows, but not for a positive definite matrix like that. It'll just go bzzz, right through. When it's tridiagonal it'll go like with the speed of light and you'll get the answer. And those five answers will be these five heights. u_1, u_2, u_3, u_4, and u_5. And we could figure it out. Actually I think section 1.2 gives the formula for this particular model problem for any size, in particular for five by five.

And there is something wonderful for this special case. The five points fall right on the correct parabola, they're exactly right. So for this particular case when the solution was a quadratic, the exact solution was a quadratic, a parabola, it will turn out-- and that quadratic matches these boundary conditions, it will turn out that those points are right on the money. So that's,

you could call, is like, super-convergence or something. I mean that won't happen every time, otherwise life would be like, too easy. It's a good life, but it's not that good as a rule. So they fall right on that curve. And we can say what those numbers are. Actually, we know what they are. Actually, I guess I could find them. What are those numbers then? And of course, one over h squared is-- What's one over h squared, just to not forget? One over h squared there, h is what? 1/6. Squared, it's going to be a 36. So if I bring it up here, bring the h squared up here, it would be times a 36. Well let me leave it here, 36. And I'm just saying that these numbers would come out right.

Maybe I'll just do the first one. What's the exact u_1, u_2? u_1 and u_2 would be what? The exact u_1, ooh! Oh shoot, I've got to figure it out. If I plug in x=1/6... Do we want to do this? Plug in x=1/6? No, we don't. We don't. We've got something better to do with our lives. But if we put that number in, whatever the heck it is, in this one, we would find out came out right. The fact that it comes out right is important.

But I'd like to move on to a similar problem. But this one is going to be free-fixed. So if this problem was like having an elastic bar hanging under its own weight and these would be the displacements points on the bar and fixed at the ends, now I'm freeing up the top end. I'm not making u_0 zero anymore. I better maybe use a different blackboard because that's so important that I don't want to erase it. So let me take the same problem, uniform bar, uniform load, but I'm going to fix u(1), that's fixed, but I'm going to free this end. And from a differential equation point of view, that means I'm going to set the slope at zero to be zero. u'(0)=0. That's going to have a different solution. Change the boundary conditions is going to change the answer. Let's find the solution. So here's another differential equation. Same equation, different boundary conditions, so how do we go? Well I had the general solution over there. It still works, right? u(x) is still -1/2 of x squared. The particular solution that gives me the one. Plus the Cx plus D that gives me zero and zero for second derivatives but gives me the possibility to satisfy the two boundary conditions. And now again, plug in the boundary conditions to find C and D. Slope is 0 at 0. What does that tell me? I have to plug that in. Here's my solution, I have to take its derivative and set x to zero. So its derivative is a 2x or a minus x or something which is zero. The derivative of that is C and the derivative of that is zero. What am I learning from that left, the free boundary condition? C is zero, right? C is zero because the slope here is C and it's supposed to be zero. So C is zero.

Now the other boundary condition. Plug in x=1. I want to get the answer zero. The answer I do

get is -1/2 at x=1, plus D. So what is D then? What's D? Let me raise that. What do I learn about D? It's 1/2. I need 1/2. So the answer is -1/2 of x squared plus 1/2. Not 1/2 x as it was over there, but 1/2. And now let's graph it. Always pays to graph these things between x equals zero and one. What does this looks like? It starts at 1/2, right? At x=0. And it's a parabola, right? It's a parabola. And I know it goes through this point. What else do I know? Slope starts at? The slope starts at zero. The other, the boundary condition, the free condition at the left-hand end, so slope starts at zero, so the parabola comes down like that. It's like half a-- where that was a symmetric bit of a parabola, this is just half of it. The slope is zero. And so that's a graph of $u(x)$.

Now I'm ready to replace it by a difference equation. So what'll be the difference equation? It'll be the same equation for the $-u''$. No change. So minus $u_{(i+1)}$ minus $2u_i$, minus $u_{(i-1)}$ over h squared equals-- I'm taking $f(x)$ to be one, so let's stay with one. Okay, big moment. What boundary conditions? What boundary conditions? Well, this guy is pretty clear. That says $u_{(n+1)}$ is zero. What do I do for zero slope? What do I do for a zero slope? Okay, let me suggest one possibility. It's not the greatest, but one possibility for zero slope is $(u_1-u_0)/h$ -- that's the approximate slope -- should be zero. So that's my choice for the left-hand boundary condition. It says u_1 is u_0. It says that u_1 is u_0.

So now I've got again five equations for five unknowns, u_1, u_2, u_3, u_4, and u_5. I'll write down what they are. Well, you know what they are. So this thing divided by h squared is all ones, just like before. And of course all these rows are not changed. But the first row is changed because we have a new boundary condition at the left end. And it's this. So u_1, well u_0 isn't in the picture, but previously what happened to u_0, when i is one, I'm in the first equation here, that's where I'm looking. i is one. It had a u_0. Gone. In this case it's not gone. u_0 comes back in, u_0 is u_1.

That might-- Ooh! Don't let me do this wrong. Ah! Don't let me do it worse! All right. There we go. Good. Okay. Please, last time I videotaped lecture 10 had to fix up lecture 9, because I don't go in. Professor Lewin in the physics lectures, he cheats, doesn't cheat, but he goes into the lectures afterwards and fixes them. But you get exactly what it looks like. So now it's fixed, I hope. But don't let me screw up. So now, what's on this top row? When i is one. I have minus u_2, that's fine. I have $2u_1$ as before, but now I have a minus u_1 because u_0 and u_1 are the same. So I just have a one in there. That's our matrix that we called T. The top row is changed, the top row is free. This is the equation T times u divided by h squared is the right-

hand side, ones. ones(5), I would call that. Properly I would call it ones(5, 1), because the MATLAB command 'ones' wants matrix and it's a matrix with five rows, one column. But it's T, that's the important thing. And would you like to guess what the solution looks like? In particular, is it again exactly right? Is it right on the money? Or if not, why not? The computer will tell us, of course. It will tell us whether we get agreement with this. This is the exact solution here and this is the exact parabola starting with zero slope. So but I solved this problem. Oh, let me see, I didn't get u_1, u_2 to u_5 in there. So it didn't look right. u_1, u_2, u_3, u_4, and u_5. And that's the right-hand side. Sorry about that. So that's T divided by h squared. T, with that top row changed, times U is the right-hand side.

By the way, I better just say what was the reason that we came out exactly right on this problem? Would we come out exactly right if it was some general load $f(x)$? No. Finite differences can't do miracles. They have no way to know what's happening to $f(x)$ between the mesh points, right? If I took this to be $f(x)$ and took this at the five points, at these five points, this wouldn't know what $f(x)$ is in between, couldn't be exactly right. It's exactly right in this lucky special case because, of course, it has the right ones. But also because, the reason it's exactly right is that second differences of quadratics are exactly right. That's what we checked on this board that's underneath there. Second differences of squares came out perfectly. And that's why the second differences of this guy give the right answer, so that guy is the answer to both the differential and the difference equation. I had to say that word about why was that exactly right. It was exactly right because second differences of squares are exactly right.

Now, again, we have second differences of squares. So you could say exactly right or no? What are you betting? How many think, yeah, it's going to come out on the parabola? Nobody. Right. Everybody thinks there's something going to miss here. And why? Why am I going to miss something? Yes? It's a first-order approximation at the left boundary. Exactly right, exactly right. It's a first-order approximation to take this and I'm not going to get it right. That first-order approximation, that error of size h is going to penetrate over the whole interval. It'll be biggest here. Actually I think it turns out, and the book has a graph, I think it comes out wrong by 1/2 h there. 1/2 h, first order. And then it, of course, it's discrete and of course it's straight across because that's the boundary condition, right? And then it starts down, it gets sort of closer, closer, closer and gets, of course, that's right at the end. But there's an error. The difference between u, the true u, and the computed u is of order h. So you could say alright, if h is small I can live with that. But as I said in the end you really want to get second order accuracy if you can. And in a simple problem like this we should be able to do it. What

I've done already covers section 1.2 but then there's a note, a worked example at the end of 1.2 that tells you how to upgrade to second order.

And maybe we've got a moment to see how would we do it. What would you suggest that I do differently? I'll get a different matrix. I'll get a different discrete problem. But that'll be okay. I can solve that just as well. And what shall I replace that by? because that was the guilty party, as you said. That was guilty. That's only a first order approximation to zero slope at zero. A couple of ways we could go. This is a correct second order approximation at what mesh point? That is a correct second order approximation to u'=0, but not at that point or at that point where? Halfway between. If I was looking at a point halfway between that, that would be centered there, that would be a centered difference and it would be good. But we're not looking there. So I'm looking here. So what do you suggest I do? Well I've got to center it. Essentially I'm going to use u_(-1). I'm going to use u_(-1). And let me just say what the effect is. You remember we started with the usual second difference here, 2, -1, -1. This is what got chopped off for the fixed method. It got brought back here by our first order method. Our second order method will-- You see what's likely to happen? That minus one is going to show up where? Over here. To center it around zero. So that guy will make this into a minus two. Now that matrix is still fine. It's not one of our special matrices.

When I say fine, it's not beautiful is it? It's got one, like, flaw, it needs what do you call it when you have your face-- cosmetic surgery or something. It needs a small improvement. So what's the matter with it? It's not symmetric. It's not symmetric and a person isn't happy with a un-symmetric problem, approximation to a perfectly symmetric thing. So I could just divide that row by two. If I divide that row by 2, which you won't mind if I do that, make that one, minus one and makes this 1/2. I divided the first equation by two. Look in the notes in the text if you can. And the result is now it's right on. It's exactly on. Because again, the solution, the true solution is squares. This is now second order and we'll get it exactly right. And I say all this for two reasons. One is to emphasize again that the boundary conditions are critical and that they penetrate into the region. The second reason for my saying this is looking forward way into October. So let me just say the finite element method, which you may know a little about, you may have heard about, it's another-- this was finite differences. Courses starting with finite differences, because that's the most direct way. You just go for it. You've got derivatives, you replace them by differences. But another approach which turns out to be great for big codes and also turns out to be great for making, for keeping the properties of the problem, the finite element method, you'll see it, it's weeks away, but when it comes, notice, the finite element

method automatically produces that first equation. Automatically gets it right. So that's pretty special. And so, the finite element method just has, it produces that second order accuracy that we didn't get automatically for finite differences.

Okay, questions on today or on the homework. So the homework is really wide open. It's really just a chance to start to see. I mean, the real homework is read those two sections of the book to capture what these two lectures have done. So Monday I'll see. We'll do elimination. We'll solve these equations quickly and then move on to the inverse matrix. More understanding of these problems. Thanks.