

The topic for today is -- Today we're going to talk, I'm postponing the linear equations to next time. Instead, I think it's a good idea, since in real life, most of the differential equations are solved by numerical methods to introduce you to those right away. Even when you see the computer where you saw the computer screen, the solutions being drawn. Of course, what really was happening was that the computer was calculating the solutions numerically, and plotting the points.

So, this is the main way, numerically, it is the main way differential equations are actually solved, if they are of any complexity at all.

So, the problem is, that initial value problem, let's write up the first-order problem the way we talked about it on Wednesday. And now, I'll specifically add to that, the starting point that you used when you did the computer experiments. And, I'll write the starting point this way. So,  $y_0$  of  $x_0$  should be  $y_0$ .

So, this is the initial condition, and this is the first-order differential equation. And, as you know, the two of them together are called an IVP, an initial value problem, which means two things, the differential equation and the initial value that you want to start the solution at. Okay, now, the method we are going to talk about, the basic method of which many others are merely refinements in one way or another, is called Euler's method. Euler, who did, of course, everything in analysis, as far as I know, didn't actually use it to compute solutions of differential equations. His interest was theoretical.

He used it as a method of proving the existence theorem, proving that solutions existed. But, nowadays, it's used to calculate the solutions numerically.

And, the method is very simple to describe.

It's so naive that you probably think that if you that living 300 years ago, you would have discovered it and covered yourself with glory for all eternity.

So, here is our starting point,  $(x_0, y_0)$ .

Now, what information do we have?

At that point all we have is the little line element, whose slope is given by  $f$  of  $(x, y)$ .

So, if I start the solution, the only way the solution could possibly go would be to start off in that direction, since I have no other information. At least it has the correct direction at  $(x_0, y_0)$ .

But, of course, it's not likely to have the correct direction anywhere else. Now, what you do, then, is choose a step size. I'll try just two steps of the method. That's, I think, good enough. Choose a uniform step size, which is usually called  $h$ . And, you continue that solution until you get to the next point, which will be  $x_0 + h$ , as I've drawn it on the picture.

So, we get to here. We stop at that point, and now you recalculate what the line element is here.

Suppose, here, the line element, now, through this point goes like that.

Well, then, that's the new direction that you should start out with going from here. And so, the next step of the process will carry you to here. That's two steps of Euler's method. Notice it produces a broken line approximation to the solution.

But, in fact, you only see that broken line if you are at a computer if you are looking at the computer visual, for example, whose purpose is to illustrate for you Euler's method. In actual practice, what you see is, the computer is simply calculating this point, that point, and the succession of points. And, many programs will just automatically connect those points by a smooth looking curve if that's what you prefer to see.

Well, that's all there is to the method.

What we have to do now is derive the equations for the method. Now, how are we going to do that? Well, the essence of it is how to get from the  $n$ th step to the  $n$  plus first step? So, I'm going to draw a picture just to illustrate that. So, now we are not at  $x_0$ .

But let's say we've already gotten to  $(x_n, y_n)$ . How do I take the next step?

Well, I take the line element, and it goes up like that, let's say, because the slope is this.

I'm going to call that slope  $A_{sub\ n}$ .

Of course,  $A_{sub\ n}$  is the value of the right hand side at the point  $(x_n, y_n)$ , and we will need that in the equation, -- -- but I think it will be a little bit clearer if I just give it a capital letter at this point. Now, this is the new point, and all I want to know is what are its coordinates?

Well, the  $x_{n+1}$  is there.

The  $y_{n+1}$  is here.

Clearly I should draw this triangle, complete the triangle.

This side of the triangle, the hypotenuse has slope  $A_n$ .

This side of the triangle has length  $h$ .

$h$  is the step size. Perhaps I'd better indicate that, actually put that up so that you know the word step.

It's the step size on the  $x$ -axis, how far you have to go to get from each  $x$  to the next one.

What's this? Well, if that slope has this, the slope  $A_n$ , this is  $h$ .

Then this must be  $h$  times  $A_n$ , the length of that side, in order that the ratio of the height to this width should be  $A_n$ . And, that gives us the method.

How do I get from, clearly, to get from  $x_n$  to  $x_{n+1}$ , I simply add  $h$ .

So, that's the trivial part of it.

The interesting thing is, how do I get the new  $y_{n+1}$ ? And so, the best way to write it as, that  $y_{n+1} - y_n$  divided by  $h$ , well, sorry,  $y_{n+1} - y_n$  is this line, the same as the line  $h$  times  $A_n$ .

So, that's the way to write it. Or, since the computer is interested in calculating  $y_{n+1}$  itself, put this on the other side. You take the old  $y_n$ , the previous one, and to it, you add  $h$  times  $A_n$ .

And, what, pray tell, is  $A_n$ ? Well, the computer has to be told that  $A_n$  is the value of  $f$ . So, now, with that, let's actually write the Euler program, not the program, but the Euler-- the Euler method equations, let's just call it the Euler equations.

What will they be? First of all, the new  $x$  is the old  $x$  plus  $h$ .

The new  $y$  is just what I've written there, the old  $y$  plus  $h$  times a certain number,  $A_n$ , and finally,  $A_n$  has the value-- It's the slope of the line element here, and therefore by definition, that's  $f$  of  $(x_n, y_n)$ .

So, if these three equations which define Euler's method.

I assume in 1.00 you must be asked to, at some point, as an exercise in the term at one point to program the computer in C or whatever they're using, Java, now, I guess, to do Euler's method.

And, these would be the recursive equations that you would put in to do that. Okay, let's try an example, then. So, what would be a good color for Euler? Well, a purple.

I assume nobody can see purple. Is that correct?

Can anyone in the back of the room see that that's purple?

Okay. Sit closer.

So, let's calculate. The example, I'll use a simple example, but it's not entirely trivial.

My example is going to be the equation,  $x^2 - y^2$  on the right hand side.

And, let's start with  $y(0) = 1$ , let's say. And so, this is my initial value problem, that pair of equations.

And, I have to specify a step size.

So, let's take the step size to 0.1.

You choose the step size, or the computer does.

We'll have to talk about that in a few minutes.

Now, what do you do? Well, I say this is a nontrivial equation because this equation, as far as I know, cannot be solved in terms of elementary functions.

So, this equation would be, in fact, a very good candidate for a numerical method like Euler's.

And, you had to use it, or maybe it was the other way around, I forget. On your problem set, you drew a picture of the direction field and answered some questions about the isoclines, how the solutions behave. Now, the main thing I want you to get, this is not just for Euler's, talking about Euler's equations. But in general, for the calculations you have to do in this course, it's extremely important to be systematic because if you are not systematic, you know,

if you just scribble, scribble, scribble, scribble, scribble, you can do the work, but it becomes impossible to find mistakes. You must do the work in a form in which it can be checked, which you can look over it and find, and try to see where mistakes are if, in fact, there are any. So, I strongly suggest, this is not a suggestion, this is a command, that you make a little table to do Euler's method by hand, I'd only ask you for a step or two, but since I'm just trying to make sure you have some idea of these equations and where they come from. So, first, the value of  $n$ , then the value of  $x_n$ , then the value of the  $y_n$ , and then, a couple of more columns which tell you how to do the calculation. You are going to need the value of the slope, and it's probably a good idea, also, because otherwise you'll forget it, to put in  $h \Delta_n$  because that occurs in the formula.

All right, let's start doing it.

The first value of  $n$  is zero. That's the starting point.

At the starting point,  $(x_0, y_0)$ ,  $x$  has the value zero, and  $y$  has the value one, so, zero and one.

In other words, starting, I'm carrying out exactly what I drew pictorially only now I'm doing it arithmetically using a table and substituting into the formulas.

Okay, the next thing we have to calculate is  $\Delta_n$ .

Well, since  $\Delta_n$  is the value of the right hand side, at the point zero one, you have to plug that in.

The right-hand side is  $x$  squared minus  $y$  squared.

So, it's  $0$  squared minus  $1$  squared.

The value of the slope, there, is minus one, negative one.

Now, I have to multiply that by  $H$ .

$h$  is  $0.1$ . So, it's negative, I'll never learn that. The way you learn to talk in kindergarten is the way you learn to talk the rest of your life, unfortunately. In kindergarten, we said minus.

Negative  $0.1$ .

$n$  is one now. What's the value of  $x_n$ ?

Well, to the old value I add  $1/10$ .

What's the value of  $y$ ? Well, at this point, you have to do the calculation. It's the old value of  $y$ .

To get this new value, it's the old value plus this number. Well, that's this plus that number is nine-tenths. An, now I have to calculate the new slope at this point. Okay, that is one-tenth squared minus nine-tenths squared. That's  $0.01$  minus  $0.81$ , which makes minus  $0.80$ , I hope.

Check it on your calculators. Whip them out and press the buttons. I now multiply that by  $h$ , which means it's going to be minus  $0.08$ , perhaps with a zero after. I didn't tell you how many decimal places. Let's carry it out to two decimal places. I think that will be good enough. And finally, the last step,  $2$ , here, add another one-tenth, so the value of  $x$  is now two-tenths.

And finally, what's the value of  $y$ ?

Well, I didn't tell you where to stop.

Let's stop at  $y$  of  $0.2$  because there's no more room on the blackboard. About approximately how big is that? In other words, this is, then, this is going to be the old  $y$  plus this number, which seems to be  $0.82$  to me.

So, the answer is, the new value is  $0.82$ .

Okay, we got a number. We did what we are supposed to do. We got a number.

Next question? Now, let's ask a few questions.

One of the first, most basic things is, you know, how right is this? How can I answer such a question if I have no explicit formula for that solution?

That's the basic problem with numerical calculation.

In other words, I have to wander around in the dark to some extent, and yet have some idea when I've arrived at the place that I want to go.

Well, the first question I'd like to answer, is this too high or too low? Is Euler, sorry, he'll forgive me in heaven, I will use him.

By this, I mean, is the result, let me just say something first, and that I'll criticize it. Is Euler too high or too low?

In other words, is the result of using Euler's method, i.e. is this number too high or too low? Is it higher than the right answer, what it should be? Or, is it lower than the right answer? Or, God forbid, is it exactly right? Well, it's almost never exactly right. That's not an option.

Now, how will we answer that question?

Well, let's answer it geometrically.

Basically, if the solution were a straight line, then the Euler method would be exactly right all the time.

But, it's not a line. Then it's a curve.

Well, the critical question is, is it curved?

Is the solution? So, here's a solution.

Let's call it  $y_1$  of  $x$ , and let's say here was the starting point. Here, the solution is convex.

And, here the solution is concave, right?

Concave up or concave down, if you learn those words, but I think those have, by now, I hope pretty well disappeared from the curriculum. Call it, if you haven't up until now, what mathematicians call it, convex is that, and the other one is concave. Well, how do Euler's solutions look? Well, I'll just sketch.

I think from this you can see already, when you start out on the Euler's solution, it's going to go like that.

Now you are too low. Well, let's suppose after that, the line element here is approximately the same as what it is there, or roughly parallel.

After all, they are not too far apart.

And, the direction field is continuous.

That is, the directions don't change drastically from one point to another. But now, you see it's still too low. It's even lower as it pathetically tries to follow. It's losing territory, and that's basically because the curve is convex.

Exactly the opposite what would happen if the curve were concave, if the solution curve were concave.

Now it's too high, and it's not going to be able to correct that as long as the solution curve stays concave.

Well, that's probably too optimistic.

It's probably more like this. So, in other words, in this case, if the curve is convex, Euler is going to be too high, sorry, too low.

Let's put E for Euler. How about that?

Euler is too low. If it's concave, then Euler is too high. Okay, that's great.

There's just one little problem left, namely, if we don't have a formula for the solution, and we don't have a computer that's busy drawing the picture for us, in which case we wouldn't need any of this anyway, how are we supposed to tell if it's convex or concave?

Back to calculus. Calculus to the rescue!

When is a curve convex? A curve is convex if its second derivative is positive because the first to be convex means the first derivative is increasing all the time.

And therefore, the second derivative, which is the derivative of the first derivative, should be positive. Just the opposite here; the curve, the slope is, the first derivative, is decreasing all the time and therefore the second derivative is negative. So, all we have to do is decide what the second derivative of the solution is.

We should probably call it a solution.

$y$  of  $x$  is a little too vague.

$y_1$  means the solution started at this point.

So, in fact, probably it would have been better from the beginning to call that  $y_1$ , except there's no room,  $y_1$ , let's say.

That means the solution which started out at the point,  $(0, 1)$ . So, I'm still talking about at a solution like that. All right, so I want to know if this is positive, the second derivative is positive

at the starting point, zero, or it's negative.

Now, again, how you can regulate the second derivative, if you don't know what the solution is explicitly, then the answer is you can do it from the differential equation itself. How do I do that?

**Well:** easy.  $y'$  equals  $x^2 - y^2$ . Okay, that tells me how to calculate  $y'$  if I know the value of  $x$  and  $y$ , in other words, the 0.01.

What would be the value of  $y''$ ?

Well, differentiate the equation.

It's  $2x - 2yy'$ .

Don't forget to use the chain rule.

So, if I want to calculate at  $(0, 1)$ , in other words, if my starting point is that curve convex or concave, well, let's calculate.  $y'$  of zero equals one.

Okay, what's  $y''$  of zero?

Well, I don't have to repeat that calculation. Using this, I've already calculated that it was negative one.

And now, the new thing, what's  $y''$  of zero?

Well, it is this.

I'll write it out. It's  $2 \cdot 0 - 2 \cdot 1 \cdot (-1)$ , which is one,  $2 \cdot 1 - 2 \cdot 0 \cdot (-1)$ , which is two.

You want to see we are pulling ourselves up by our own boot straps, which is impossible.

But, it is not impossible because we are doing it.

So, what's the answer? Zero here, two, I've calculated without having the foggiest idea of what the solution is or how it looks. I've calculated that its second derivative at the starting point is two.

Therefore, my solution is convex at the starting point.

And therefore, this Euler approximation, if I don't carry it out too far, will be too low.

So, it's convex Euler, too low.

Now, you could argue, yeah, well, what about this?

[LAUGHTER] So, you could go like this, and then you can see it catches up.

Well, of course, if the curve changes from convex to concave, then it's really impossible to make any prediction at all. That's a difficulty.

So, all this analysis is only if you stay very nearby.

However, I wanted to show you, the main purpose of it in my mind was to show you how do you use, it's these equations, how to use the differential equation itself to get information about the solutions, without actually being able to calculate the solutions? Now, so that's the method, and that's how to find out something about it.

And now, what I'd like to talk about is errors.

How do I handle, right?

So, in a sense, I've started the error analysis. In other words, the error, by definition, the error is this difference,  $e$ .

So, in other words, what I'm asking here, is the error positive?

It depends which we measure it. Usually, you take this minus that. So, here, the error would be considered positive, and here it would be considered negative, although I'm sure there's a book somewhere in the world, which does the opposite. Most hedge by just using the absolute value of the error plus a statement that the method is producing answers which are too low or too high.

The question, then, is, naturally, this is not the world's best method.

It's not as bad as it seems. It's not the world's best method because that convexity and concavity means that you are automatically introducing a systematic error.

If you can predict which way the error is going to be by just knowing whether the curve is convex or concave, it's not what you want. I mean, you want to at least have a chance of

getting the right answer, whereas this is telling you you're definitely going to get the wrong answer. All it tells you is, and it's telling you whether your answer is going to be too high or too low. We've like a better chance of getting the right answer. Now, so the question is, how do you get a better method? A search is for a better method. Now, the first method, which will occur, I'm sure, to anyone who looks at that picture, is, look, if you want this yellow line to follow the white one, the white solution, more accurately, for heaven's sake, don't take such big steps. Take small steps, and then it will follow better. All right, let's draw a picture. Excuse me.

My little box of treasures, here.

[LAUGHTER] So, use a smaller step size.

And the picture, roughly, which is going to justify that, will look like this.

If the solution curve looks like this, then with a big step size, I'm liable to have something that looks like that.

But, if I take a smaller step size, suppose I halve the step size. How's it going to look, then? Well, I better switch to a different color. If I halve the step size, I'll get a littler, goes like that.

And now it's following closer. Of course, I'm stacking the deck, but see how close it follows?

I'm definitely not to be trusted on this.

Okay, let's do the opposite, make really big steps.

Suppose instead of the yellow ones I used the green one of double step size. Well, what would have happened then? Well, I've started out, but now I've gone all the way to there.

And now, on my way up, of course, it has a little further to go. But, if for some reason, I stop there, you could see, I would be still lower. In other words, the bigger the steps size, the more the error.

And, where are the errors that we are talking about?

Well, the way to think of the errors, this is the error, that number the error. You can make it positive, negative, or just put it automatically an absolute value sign around it. That's not so important.

So, in other words, the conclusion is, that the error  $e$ , the difference between the true value that I should have gotten, and the Euler value that the calculation produced, that the error  $e$ , depends on the step size. Now, how does it depend on the step size? Well, it's impossible to give an exact formula, but there's an approximate answer, which is, by and large, true. So, the answer is,  $e$  is going to be a function of  $h$ .

What function? Well, asymptotically, which is another way of putting quotation marks around, what did I say? It's going to be a constant, some constant, times  $H$ .

[LAUGHTER] It looks like this, and for this reason it's called a first order, the Euler is a first-order method.

And now, first-order does not refer to the first order of the differential equation. It's not the first order because it's  $y' = f(x, y)$ .

The first order means the fact that  $h$  occurs to the first power.

The way people usually say this is since the normal way of decreasing the step size, as you'll see as is you try to use a computer visual that deals with the Euler method, which I highly recommend, by the way, so highly recommended that you have to do it, is that the way to say it, each new step halves the step size.

That's the usual way to do it. If you halve the step size, since this is a constant, if I halve the step size, I halve the error, approximately.

Halve the step size, halve the error.

That tells you how the error varies with step size for Euler's method. Please understand, that's what people say, and please understand the grammatical construction. Since everyone in the math department has a cold these days except me for the moment, everyone goes around chanting this mantra.

This is totally irrelevant. This whole mantra, feed a cold, starve a fever.

And if you asked them what it means, they say eat a lot if you have a cold. And if you have a fever, don't eat very much, which is not what it means at all. Grammatically, it's exactly the same construction as this.

What this means is if you halve the step size, you will halve the error. That's what feed a cold, starve a fever means. And, remember this for the rest of your life. If you feed a cold, if you eat too much when you have a cold, you will get a fever and end up still having to starve yourself because, of course, nobody, when you have a fever, nobody feels like eating, so they don't eat anything.

All right, you got that? Good.

I want all of you to go home and tell that to your mothers.

You know, that's the way we always used to speak.

**Grimmer ones:** spare the rod, spoil the child does not mean that you should not hit your kid. It means that if you fail to hit your kid, he or she will be spoiled, whatever that means. So, you don't want to do that.

I guess the mantra today would be, I don't know.

Okay, so the first line of defense is simply to keep having the step size in Euler. And, what people do is, if they don't want to use anything better than Euler's method, is you keep having the step size until the curve doesn't seem to change anymore. And then you say, well, that must be the solution.

And, I asked you on the problems set, how much would you continue to have to halve the step size in order for that good thing to happen?

However, there are more efficient methods which get the results faster. So if that's our good method, let's call this our still better method.

The better methods aim at being better.

They keep the same idea as Euler's method, but they say, look, let's try to improve that slope, An. In other words, since the slope, An, that we start with is guaranteed to be wrong if the curve is convex or concave, can we somehow correct it? So, for example, instead of immediately aiming there, can't we somehow aim it so that by luck, we just, at the next step just lands us back on the curve again?

In other words, with sort of looking for the short path, a shortcut path, which by good luck will end us up back on the curve again. And, all the simple improvements on the Euler method,

and they are the most stable in ways to solve differential equations numerically, aim at finding a better slope.

So, they find a better value for a better slope, find a better value than  $A_n$ . Try to improve that slope that you found. Now, once you have the idea that you should look for a better slope, it's not very difficult to see what, in fact, you should try. Again, I think most of you would say, hey, I would have thought of that.

And, you would be closer in time, since these methods were only found about around the turn of the last century is when I place them, mostly by some German mathematicians interested in solving equations numerically.

All right, so what is the better method?

Our better slope, what should we look for in our better slope? Well, the simplest procedure is, once again, we are starting from there, and the Euler slope would be the same as a line element.

So, the line element looks like this.

And, our yellow slope,  $A$ , and I'll still continue to call it  $A_n$ , goes like that, gets to here.

Okay, now if it were convex, if the curve were convex, this would be too low. And therefore, the next step would be, I'm going to draw this next step in pink. Well, let's continue in here, would be going up like that. I'll call this  $B_n$ , just because it's the next step of Euler's method.

It could be called  $A_n$  prime or something like that.

But this will do. And now what you do is, let me put an arrow on it to indicate parallelness, go back to the beginning, draw this parallel to  $B_n$ .

So, here is  $B_n$ . Again, just a line of that same slope. And now, what you should use as the simplest improvement on Euler's method, is take the average of these two because that's more likely to hit the curve than  $A_n$  will, which is sure to be too low if the curve is convex. In other words, use this instead. Use that.

So, this is our better slope. Okay, what will we call that slope? We don't have to call it anything. What were the equations for the method be? Well,  $x_{n+1}$  is gotten by adding the step size.

So, here's my step size just as it was before.

Just as it was before, the new thing is how to get the new value of  $y$ . So,  $y_{n+1}$  should be the old  $y_n$ , plus  $h$  times not this crummy slope,  $A_n$ , but the better, the pink slope.

What's the formula for the pink slope?

Well, let's do it in two steps. It's the average of  $A_n$  and  $B_n$ .

Hey, but you didn't tell me, or I didn't tell you what  $B_n$  was. So, you now must tell the computer, oh yes, by the way, you remember that  $A_n$  was what it always was. The interesting thing is, what is  $B_n$ ? Well, to get  $B_n$ ,  $B_n$  is the slope of the line element at this new point.

Now, what am I going to call that new point?

I don't want to call this  $y$  value,  $y_{n+1}$ , because that's, it's this up here that's going to be the  $y_{n+1}$ . All this is, is a temporary value used to make another calculation, which will then be combined with the previous calculations to get the right value. Therefore, give it a temporary name. That point, we'll call it, it's not going to be the final, the real  $y_{n+1}$ .

We'll call it  $y_{n+1}$  twiddle,  $y_{n+1}$  temporary. And, what's the formula for it?

Well, it's just going to be what the original Euler formula; it's going to be  $y_n$  plus what you would have gotten if you had calculated, in other words, it's the point that the Euler method produced, but it's not, finally, the point that we want.

Now, do I have to say anything else?

Yeah, I didn't tell the computer what  $B_n$  was.

Okay,  $B_n$  is the slope of the direction field at the point  $n+1$ . And the computer knows what that is. And, this point,  $y_{n+1}$  temporary.

So, you make a temporary choice of this, calculate that number, and then go back, and as it were, correct that value to this value by using this better slope. Now, that's all there is to the method, except I didn't give you its name.

Well, it has three names, four names in fact.

Which one shall I give you? I don't care.

Okay, the shortest name is Heun's method.

But nobody pronounces that correctly.

So, it's Heun's method. It's called, also, the Improved Euler method.

It's called Modified Euler, very expressive word, Modified Euler's method. And, it's also called RK2.

I'm sure you'll like that name best.

It has a Star Wars sort of sound.

RK stands for Runge-Kutta, and the reason for the two is not that it uses, well, it is that it uses two slopes, but the real reason for the two is that it is a second-order method. So, that's the most important thing to put down about it. It's a second-order method, whereas Euler's was only a first-order method.

So, Heun's method, or RK2, let's write it, is the shortest thing to write, is a second-order method, meaning that the error varies with the step size like some constant, it will not be the same as the constant for Euler's method, times  $h$  squared.

That's a big saving because it now means that if you halve the step size, you're going to decrease the error by a factor of one quarter. You will quarter the error.

Now, you say, hey, why should anyone use anything else? Well, think a little second.

The real thing which determines how slowly one of these methods run is you look at the hardest step of the method and ask how long does the computer take, how many of those hardest steps are there? Now, the answer is, the hardest step is always the evaluation of the function because the functions that are common use are not  $x$  squared minus  $y$  squared. They take half a page and have, as coefficients, you know, ten decimal place numbers, whatever the engineers doing it, whatever their accuracy was. So, the thing that controls how long a method runs is how many times the slope, the function, must be evaluated.

For Euler, I only have to evaluate it once.

Here, I have to evaluate it twice.

Now, roughly speaking, the number of function evaluations will each give you the exponent.

The method that's called Runge-Kutta fourth order will require four evaluations of slope, but the accuracy will be like  $h$  to the fourth: very accurate.

You halve the step size, and it goes down by a factor of

16. Great. But you had to evaluate the slope four times.

Suppose, instead, you halve four times this thing.

And, what would you have done? You would have decreased it to  $1/16$ th to what it was. You still would increase the number of function evaluations you needed by four, and you would have decreased the error by a  $16$ th.

So, in some sense, it really doesn't matter whether you use a very fancy method, which requires more function evaluations. That's true.

The error goes down faster, but you are having to more work to get it. So, anyway, nothing is free.

Now, there is an RK4. I think I'll skip that, since I wouldn't dare to ask you any questions about it.

But, let me just mention it, at least, because it's the standard. It uses four evaluations.

It's the standard method, if you don't want to do anything fancier. It's rather inefficient, but it's very accurate, standard method, accurate, and you'll see when you use the programs, it's, in fact, a program which is drawing those curves, the numerical method which draws all those curves that you believe in on the computer screen is the RK4 method. The Runge-Kutta, I should give them their names. Runge-Kutta, fourth-order method. Two mathematicians, I believe both German mathematicians around the turn of the last century, Runge-Kutta fourth-order method requires four slopes, requires you to calculate four slopes. I won't bother telling you what you do, but it's a procedure like that.

It's just a little bit more elaborate.

And you take two of these, you make up a weighted average for the super slope. You use weighted average.

What should I divide that by to get the right...?

Six.

Why six? Well, because if all these numbers were the same, I'd want it to come out to be whatever that common value was. Therefore, in a weighted average, you must always divide by the sum of the coefficients.

So, this is the super-slope. And, if you plug that super-slope into here, you will be using the Runge-Kutta method, and get the best possible results. Now, I wanted to spend the last three minutes talking about pitfalls of numerical computation in general. One pitfall I am leaving you on the homework to discover for yourself.

Don't worry, it won't cause you any grief.

It'll just destroy your faith in these things for the rest of your life, which is probably a good thing.

So, pitfalls, number one, you find, you'll find. Let me talk, instead, briefly about number two, which I am not giving you an exercise in. Number two is illustrated by the following equation. What could be simpler?

This is a very bad equation to try to solve numerically.

Now, why? Well, because if I separate variables, why don't I save a little time?

I'll just tell you what the solution is, okay?

You obviously separate variables.

Maybe you can do it in your head.

The solution will be, the solutions will have an arbitrary constant in them, and they won't be very complicated. They will be one over c minus x.

C is an arbitrary constant, and as you give different values, you get, now, what do those guys look like?

Okay, so here I am. I start out at the point, one. And, I start out, I tell the computer, compute

for me the value of the solution at one starting out at one.

And, it computes and computes a little while.

But the solution, how does this curve actually look? So, in other words, suppose I say that  $y$  of zero equals one.

Find me  $y$  of two. In other words, take a nice small step size. Use the Runge-Kutta fourth-order method. Calculate a little bit, and tell me, I just want to know what  $y$  of two is. Well, what is  $y$  of two?

Well, unfortunately, how does that curve look?

The curve looks like this. At that point, it drops to infinity in a manner of speaking, and then sort of comes back up again like that.

What is the value of  $y$ ? This is the point, one. What is the value of  $y$  of two?

Is it here?

Is it this? Well, I don't know, but I do know that the computer will not find it.

The computer will follow this along, and get lost in eternity, in infinity, and see no reason whatever why it should start again on this branch of the curve.

Okay, well, can't we predict that that's going to happen somehow, avoid what I should have.

The whole difficulty is, this is called a singular point. The solution has a singularity, in other words, a single place where it goes to infinity or becomes discontinuous, maybe as a jump discontinuity. It has a singularity at  $x$  equals  $c$ . This, in particular, at  $x$  equals one here, but from the differential equation, where is that  $c$ ? There is no way of predicting it. Each solution, in other words, to this differential equation, has its own, private singularity, which only it knows about, and where it's going to blow up, and there's no way of telling from the differential equation where that's going to be.

That's one of the things that makes numerical calculation difficult, when you cannot predict where things are going to go bad in advance.