HERBERT GROSS: Hi. Today we begin our discussion of differential equations. It has been said that the language of the universe is differential equations. And for this reason, I would suspect from a practical point of view differential equations is most important.

For example, one is used to stating physical principles in terms of rates of change. To say, for example, that the rate of change is proportional to the amount of present is a rather simple sounding physical principle. The idea is knowing that the rate of change is proportional to the amount present.

We sometimes like to find out explicitly what the amount looks like as a function of time. And that's what we mean by a solution or a general solution to differential equations. And that will be our subject for our first lesson in our study of differential equations-- the concept of a general solution.

The first and simplest type of equation that we're going to talk about is the differential equation which we say has order 1. In other words, the greatest derivative that appears is the first derivative. There are no second derivatives involved-- in other words, a problem where, in terms of cannot kinematics, you are told what the velocity looks like in terms of the time-- in other words-- and you want to find the answer in terms of displacement. dv dt equals-- dx dt equals something. Find what x looks like explicitly in terms of t.

Now, in general, what we're saying is all we know about a first order differential equation is that there is some relationship between the independent variable x, the dependent variable y, and dy dx. And the symbolic way of writing that is to say that some function-- say, capital F-- of x, y, and dy dx is 0. That's just a mathematical symbolism for writing down the most general first order differential equation.

Now the question that comes up is twofold. One is, first of all, does equation 1 even have a solution? How do we know that just because we can state the differential equation that there is a specific function-- say, y as a function of x-- that satisfies the equation? And the second is, assuming that there is a function y as a function of x-- that y as a function of x does satisfy this equation, how do we know that that solution is unique?

Two questions. Does the equation have a solution? Is the solution unique? And I think that the easiest way to illustrate this is in terms of part one of our course, where we already solve a rather simple type of differential equation, namely that differential equation that had the specific form dy dx is some function of x.

In particular, suppose I were to say find all solutions of dy dx equals 2x-- and I'll just add on here-- which pass through a particular point x0, y0. I'll come back to that in a moment. By the way, notice what I mean by an explicit solution. If somebody says to me, find all functions y such that dy dx equals 2x, in the language of sets, notice that somebody could get kind of cute with us and say, gee, isn't that just a set of all functions y such that dy dx equals 2x?

And the answer is, that is correct. Of course, that's the answer. But what we mean is we would like to know what y looks like explicitly in terms of x. In other words, given x, what is y? This only tells me what dy d looks like explicitly in terms of x. So what I am really interested in in problems of this type is to convert from this set builder-- this implicit solution set-- to an explicit solution set.

And as a trivial review, recall that, in this case, we knew that if dy dx was 2x then y, had to be x squared plus c. In other words, the solution set S written in explicit form is the set of all y such that y equals x squared plus c. And I suppose if you wanted a definition of what it means to solve a differential equation, why couldn't we just say-- just like we do an algebra in a sense-- that solving a differential equation essentially means to convert from the set builder-- the implicit form-- to the explicit form.

So in other words, explicitly to express y in terms of x, the set of all y such that y equals x squared plus c. So now I know what y looks like in terms of x, up to an arbitrary constant. It is the solution set of the differential equation dy dx equals 2x.

Now you may have noticed I said it as part of the problem, let's see if there is a solution which passes to a particular point x0, y0. What I do is as I look at this particular equation and I replace x by x0 and y by y0 and solve for see. And by the way, I would intuitively expect to be able to find that value of c. Because after all, I have an arbitrary constant in here, which means I have a degree of freedom. And therefore, I would suspect that by replacing x by x0 and y by y0, I should be able to find a specific value of c.

Now notice what this leads to if I replace x by x0, y by y0, I get y0 equals x0 squared plus c. Consequently, c is y0 minus x0 squared. x0 and y0 were specific numbers, real numbers. Consequently, y0 minus x0 squared is a specific number. And consequently, the specific choice of c, such that the equation becomes what y equals x squared plus the constant y0 minus x0 squared, is the only solution of what the equation dy dx equals 2x that passes through the point x0, y0.

This is kind of baby stuff in terms of the fact that it was elementary in part 1 of our course. But remember, this solution set is what? It's a set of parabolas which are, essentially, a parallel family of parabolas. And what you're sort of saying is, gee, it seems obvious, geometrically, that since I have this infinite family of parallel parabolas that's specifying a point that has to be on the parabola uniquely determines a member of that family.

At any rate, I don't want to belabor this point. I want to get to the greater subtleties that come up in handling first order differential equations. So I've picked a slightly tougher problem for example 2. Example 2-- and by the way, this is a special equation I've picked. It's known as Clairaut's Equation. And the equation is this. It's y equals x dy dx minus dy dx squared.

Clairaut's Equation is any equation which has the form y equals x times dy dx plus something which depends only on dy dx. In other words, rather than write this in an abstract form, I decided to pick a particular application of this. And that is, let's just pick this particular example. That's one special form of Clairaut's Equation. It's y equals x dy dx minus dy dx squared. x times dy dx, and what's left is a function of dy dx alone.

At any rate, it turns out that one explicit solution of Clairaut's Equation-- it's a very interesting thing-- is you simply replace dy dx by c. Very amazing thing. In other words, if you have a Clairaut's Equation, and you want a solution, every place you see dy dx, replace it by c. For example, in this case, I would get what? y is equal to xc, or cx-- I guess you like to write the constant first-- cx minus c squared, which, by the way, is what a straight line whose slope is c and whose y-intercept is minus c squared.

By the way, the proof that this is a solution is simply differentiate this. You get that dy dx equals c. Since c is dy dx, every place you see a c, replace it by dy dx. And that will give you what? y equals x dy dx minus dy dx squared. I'll give you more on that in the exercises on how one solves Clairaut's Equations. I want it for illustrative purposes here.

So anyway, this is a solution of this particular equation. And it's only one solution. I don't know if there are other solutions, or what have you. The next question that I would investigate here, just as in the previous case-- I would say, gee, I wonder what c must be if I want a solution curve to pass through the point x0, y0 in the xy-plane.

So what I do, purely algebraically, is I replace x by x0, y by y0, and solve the resulting equation for c-- namely y0 equals x0 c minus c squared. And I now observe-- remember, x and y look like variables. With a subscript 0, they represent fixed x and y coordinates of a point. c is the only variable in here. In other words, this is a quadratic equation in c. And solving this quadratic equation by the quadratic formula, I find that c is equal to this expression here.

Remembering that we're dealing with real variables, I now know that I'm in trouble if this thing here happens to be negative. In other words, if this is negative, I have an imaginary number when I take the square root. In other words, there will be no solution, at least of this particular type. There'll be no solution to this equation if y happens to be-- if y0 is greater than 1/4 x0 squared. And that means there will be no solution if you're above the parabola y equals 1/4 x squared.

Remember, the set of all points x comma y for which y is greater than 1/4 x squared means you're above the parabola y equals 1/4 x squared, because on that parabola, since y is equal to 1/4 x squared, you would have to go above that to make y greater than 1/4 x squared.

By the way, you see, all this proves is that there is no solution of this type, to this particular equation, if you're above the parabola y equals 1/4 x squared. The question is, what if I didn't give you this as a hint? Suppose I just gave you this equation, didn't even give you a hint as to what the solution should be, and just said, tell me where there will be solutions?

Notice that you did not need this piece of information to get started on this problem. Namely, all you had to say was, gee, this is a quadratic equation in dy dx. Let me solve this quadratic equation. See, what I could have done was what? I could have said, dy dx squared minus x dy dx plus y equals 0, and then solved that equation for dy dx.

Without boring you with the details of a quadratic formula all over again, it would have turned out, more generally, that dy dx is equal to x plus or minus the square root of x squared minus 4y over 2. And then I would have seen, right from this, without even knowing what the solution was, that dy dx isn't even a real number.

Remember, dy dx, geometrically, is slope. And slope is real. So this wouldn't even be a real number if y was greater than 1/4 x squared. So no matter what the solution is, the geometry of this problem tells me there can be no solution if y is greater than 1/4 x squared. And what that means is, if I draw the graph now-- see, in other words, this is the x-axis, the y-axis. This is the parabola y equals 1/4 x squared. There are no solutions that pass through any points in here. No solutions in here.

By the way, the plus or minus sign here is relatively unimportant. What it means is, dy dx is a double-valued function. We would solve this in real life as two separate problems-- namely, dy dx equals x plus the square root of x squared minus 4y over 2, and dy dx is x minus the square root of x squared minus 4y over 2. Notice that the plus or minus sign has no bearing on the fact that the crucial point is that y must be no greater than 1/4 x squared, all right?

By the way, if you want to see this geometrically, notice that y equals cx minus c squared is what kind of a line? It's a straight line whose y-intercept is minus c squared. Notice that minus c squared is negative whether c is positive or negative, because c squared can't be negative. So minus c squared can't be positive. Notice, however, if c is positive-- c is the slope of this line-- if c is positive, the straight line goes through this point, all right? And has this slope. If c is negative, it goes through this point with this slope.

There's a very amazing property that the parabola 1/4 x squared has with respect to this family of straight lines. And I'll also talk about that more in the exercises. The parabola y equals 1/4 x squared is called the envelope of this family of straight lines. The amazing thing is that every straight line in this family is tangent to this parabola. And conversely, the tangent line at any point of this parabola is a line that belongs to this family.

And that leads to a very remarkable thing, which is not immediately, intuitively obvious. And that is, in particular, it must mean that the parabola itself must be a solution of my differential equation. You see, after all, if I pick a point, now, that's on the parabola and draw the tangent line in here, that tangent line belongs to this family. Because this point is on this tangent line, it must be a solution to the original differential equation.

But that point is also on the parabola. The point doesn't know whether I'm considering it as being part of the parabola or a part of the family of straight lines. Consequently, what it means is that this parabola must also satisfy the differential equation. And the best way of proving that is to show you that it does satisfy the equation. Namely, if y equals 1/4 x squared, notice that dy dx is x over 2. Consequently, dy dx squared is x squared over 4.

The differential equation I'm trying to satisfy is this one. If I replace y by 1/4 x squared, dy dx by x over 2-- and look at what this says. It says x squared over 4 is equal to x squared over 2 minus x squared over 4, which, of course, is x squared over 4. And indeed, the parabola does satisfy this equation.

And now you see we've run into a very interesting situation here. Namely, on the parabola y equals 1/4 x squared, y equals d dy dx minus dy dx squared has at least two solutions-- namely, a member of the family of straight lines y equals cx minus c squared, and also the parabola y equals 1/4 x squared.

Remember, at the beginning of our lecture today, we mentioned two questions. First of all, does the differential equation have a solution? And if it does have a solution, is the solution unique? What we've shown so far is that on the parabola, the solution is not unique. Namely, there were two solutions on the parabola. And we've shown that, above the parabola, the equation has no solutions. Consequently, the answers to both questions, 1 and 2, can be in the negative.

And this is what makes differential equations a very tough subject, and why it's difficult to talk about the solution, or a solution, or what have you. But fortunately, there is a key theorem which we have at our disposal-- a theorem which is far more difficult to prove than it is to state and memorize. And I will settle right now for just stating the theorem and having you see what it means.

The key theorem is this. Let's assume, for the sake of argument, that we can write our differential equation explicitly in terms of dy dx-- in other words, that dy dx is explicitly some function of x and y. If it turns out that f and its partial with respect to y are continuous in some region R, then at each point in R, there is a unique solution of dy dx equals f of x, y which passes through x0 y0.

You see, notice that in the particular problem that we were just dealing with, what was f of x, y? f of x, y was this function here. Notice that f of x, y will be continuous as long as this expression exists. In other words, f of x, y will be continuous as long as y is less than or equal to 1/4 x squared. What about the partial of f with respect to y? To find the partial of f with respect to y, I have to differentiate this thing.

Notice that y is under the square root sign. When I differentiate the square root, the square root comes down into the denominator. I can't allow a 0 denominator. So that tells me that not only must y be less than or equal to 1/4 x squared, but rather, y must be less than 1/4 x squared, because if y equaled 1/4 x squared, when this factor comes down into the denominator, I'm in trouble. I have a 0 denominator.

Notice, by the way, the bad solution that we got. In other words, notice that where we got the two solutions occurred where y was equal to 1/4 x squared. And our key theorem doesn't apply in that case, because notice that our key theorem says that the solution will be unique only in that region R where f and f sub y are continuous.

You see, notice that, below the parabola, f and f sub y were continuous. In other words, these functions couldn't go bad in our problem. Consequently, since we had one solution of the form y equals cx minus c squared that passed through every point that was below the parabola, the fact that this theorem applies says there can't be any other solution, because once you've found one solution, that's all there are. That's exactly what you mean by unique.

See, in other words, if the solution is unique, then we can talk about a general solution-- namely, that solution that has one arbitrary constant. We can find what point it passes through by solving for the arbitrary constant. And that's what we mean by a particular solution. In fact, let me just summarize that for you in terms of our two examples.

By a general solution, you mean what? You mean a solution that has one arbitrary constant such that, through any point in your region, you can get one and only one solution. Notice that the differential equation dy dx equals 2x has one solution, y equals x squared plus c-- one family of solutions, all right? That's the general solution.

Notice that 2x-- you see, if f of xy is 2x, f is certainly continuous. The partial of f with respect to y, since f of xy is just 2x is a function of x alone, the partial with respect to y is 0, which is certainly a continuous function. In other words, that's another way of seeing what we had in part one-- another way of looking at it.

That's, in terms of our key theorem here, why we have a unique family of solutions-- that in the whole xy-plane, 2x and its partial with respect to y happen to be continuous. On the other hand, in example 2, the general solution was y equals cx minus c squared provided that our region R was restricted to being what? That we were below the parabola y equals 1/4 x squared.

By a particular solution, we mean a solution that you can get to pass through a particular point. In other words, by arbitrarily specifying a value of the constant. For example, referring to example 1 again, if I pick c to be 7 or minus pi [INAUDIBLE] y equals x squared plus 7, y equals x squared minus pi, these are particular solutions of the equation that we're talking about.

In particular, what we're saying is, if you want the particular solution-- see, maybe I should emphasize that-- the particular solution that passes through the point x0, y0, your constant must be chosen to be y0 minus x0 squared. In example 2, the general solution, if we were below the parabola, was y equals cx minus c squared. In particular, if I were to pick c to be 3, a particular solution would be what? y equals 3x minus 9.

Or to do this more generally, given the point x0, y0, the particular solution of y equals cx minus c squared, which passes through this point, is determined by c being this expression, where the plus or minus sign is not really ambiguous here. What we meant was, is that when we have the plus or minus sign, we look at this as two separate problems.

Finally, we come to a third concept that's called a singular solution. A singular solution occurs only when you're in the region where f or its partial with respect to y are not continuous. For example, there are no singular solutions to example 1. On the other hand, a singular solution to example 2 was y equals 1/4 x squared.

Why do I call it a singular solution? Let me point out that there is no way of getting y equals 1/4 x squared by judiciously choosing a constant c. As long as c is a constant, observe that, for every choice of c, y equals cx minus c squared is a straight line-- in particular, the straight line whose slope is c and whose y-intercept is minus c squared. In other words, I cannot get the curve y equals 1/4 x squared by specifying a constant here. That's why it's called a singular solution. It cannot be obtained from the general solution.

But notice that that singular solution only existed because we violated the condition that our region be below the parabola. In other words, once we get outside of the region where f and f sub y were continuous, then mongrel-type solutions could have snuck in. And those are what are called singular solutions. But I'll talk about those more.

In the exercises, what I thought I would like to do for the remainder of this lecture is to talk specifically about what the textbook is all about and what we'll be dealing with for the next few lessons. You see, the next few units will not have any lectures. This is the lecture that will launch you into solving first order, first degree equations.

From this point on, until we get to second order equations, there will be no lectures-- simply reading assignments in the texts and learning exercises, where I'll try to give you the so-called cookbook part of differential equations. But I thought that, to lead into that, what I thought I wanted to bring out was this.

Let's suppose that we restrict our attention to not only first order equations, but first degree equations. In other words, not only does the only derivative that appears is-- be the first derivative, but it also happens to appear only to the first power, all right? Notice that, in that situation, as I look at every term in my differential equation, they fall into two types. There'll be a term in which dy dx is a factor or a term where dy dx is not a factor.

What I can therefore do is collect all the terms which have dy dx as a factor and factor out dy dx. So I have a term of the form what? Some function N of x and y times dy dx plus what? A bunch of terms involving only x and y, which have no dy dx in them-- equals 0. In other words, this is the most general form of my first order, first degree equation.

If I now treat these as differentials and multiply through by dx, notice that I'm back to the standard differential equation form that we talked about-- in fact, I'll bring this up again in a moment-- in block 3 when we introduced exact differentials. Namely, I have an equation of what form? M dx plus N dy equals 0.

By the way, to correlate this with our key theorem, notice that, from here, I could have written that dy dx was minus M of x, y over N of x, y. And that's the f of x, y that I'm talking about in our key theorem that we talked about earlier. See, in other words, f of x, y is the right-hand side of the equation when the left-hand side is specifically dy dx.

So in other words, what I'm guaranteed of is this. As long as this function on the right-hand side is continuous, and its partial with respect to y is continuous, I'm guaranteed that this equation has a general solution. In particular, if M and N happen to be continuously differential functions, the general solution will exist, provided only that N and its partial with respect to y are not 0, because, after all, to differentiate this with respect to y, this is a quotient.

And the only place that I'm going to be in trouble is if the square of the denominator appears when I use the quotient rule. So I have to be careful where either the partial with N respect to y or N itself happen to be 0. But if that doesn't happen, it means there is a general solution.

Now, you may remember, back in block 3, we said, if this differential happens to be exact-- in other words, if there happens to be a function w such that dw is M dx plus N dy, then it's trivial to solve this equation. In particular, what we mean by an exact differential equation is simply this.

If M dx plus N dy is exact, then we call the equation M dx plus N dy equals 0 an exact differential equation. And the solution of that equation is simply f of x, y equals c, where f is any function such that df is M dx plus N dy. And we're sure that such a function f exists by definition of M dx plus N dy being exact.

See, by way of example, suppose I were given the equation y dx plus x dy equals 0. Forget about the fact that I could have solved this easier by just separating the variables. Notice, by way of illustration, that the left-hand side here is just the differential of x times y. In other words, y dx plus x dy equals 0 says that the differential of xy is 0, where y is implicitly, now, some function of x. Now, if the differential of some function of x is 0, then that function itself must be a constant. Consequently, xy equals a constant is a solution of this differential equation.

The problem that comes up is, wouldn't it be nice-- that's not the problem. The question is, wouldn't it be nice if every first order, first degree differential equation happened to be exact? See, if it were, we'd be all done. We talked about that, as I say, back in block 3. If this were exact, this is how we solve it. But we also saw, in block 3, that it's very unique if the thing happens to be exact.

See, in general, the differential won't be exact. And that's why I look now at non-exact equations. See, non-exact means what? Not that the equation isn't exact, but the differential isn't an exact differential. Look at y dx minus x dy. The partial of y with respect to y is 1. The partial of minus x with respect to x is minus 1. This thing is not exact. Consequently, this is not the differential of any particular function.

Let me show you a little trick over here, though. This sort of suggests the quotient rule, see? The denominator times the differential of the numerator minus the numerator times the differential of the denominator-- except there should be a denominator here appearing as y squared, you see?

And what I do is I say, OK. Why don't I just divide both sides of this equation by y squared? If I do that and invoke the rule that equals divided by equals are equal, being careful to remember that I'm in trouble when y equals 0-- I've got to keep that in the back of my mind, remembering that y is 0-- at any rate, this equation transforms into y dx minus x dy over y squared equals 0.

The left-hand side-- this differential is now exact, believe it or not. It wasn't exact to begin with. But by dividing it by y squared, it became exact. In fact, the left-hand side is now the differential of x over y. In other words, the differential of x over y is 0. Therefore, x over y is a constant. Or, in particular, y is some constant times x.

By the way, remembering that y had to be unequal to 0 for this to be true, notice that we have to check y equals 0 separately. Notice, by the way, that y equals 0 is a particular solution of this equation obtained by choosing c equal to 0. So y equals 0 gives us no great hardship here.

But the thing I'm leading up to-- and I am sorry for all these asides-- but what I'm leading up to is something called an integrating factor. And that's what we just used over here. What we say is, if M dx plus N dy is not 0-- equals 0-- is not exact, make it exact. What does that mean? Find a function u of x and y such that when I multiply or divide-- it makes no difference-- both sides of this equation by it, the new equation, which you see, has the same solution set as this one.

See, in other words, something is a solution of this equation if and only if it's a solution of this equation. See, it's, equals multiplied by equals are equal. What if this is exact? See, that's precisely what we did up here. y dx minus x dy was not exact. The u of xy, in that case, was just 1 over y squared. We multiplied both sides of the equation by 1 over y squared and made this exact.

The problem is this, though. It turns out, theoretically, that if an equation of first order, first degree is not exact, there is an integrating factor-- in other words, something that you can multiply it by to make it exact. The problem is that, in real life, that factor is harder to find than the solution of the original equation, in most cases. Namely, how would I find a u such that this was exact?

Remember, the condition for exactness is that the partial of this with respect to y has to equal the partial of this with respect to x. Remembering that u is a function of x and y, I have to use the product rule over here. I get u times the partial of M with respect to y plus the partial of u with respect to y times M equals u times the partial of N with respect to x plus the partial of u with respect to x times N.

And look at this equation. I'm trying to solve this for u. And notice that this is still a differential equation, but it involves even partial derivatives. In other words, to solve this equation involves having to solve a partial differential equation, which is even harder to do than the equation that we're trying to solve now.

Well, at any rate, let me show you now, in summary, why people refer to differential equations as a cookbook course. Philosophically, there's nothing to solving first order, first degree differential equations. Namely, look at the differential. If it's exact, bang. You just integrate it directly. Find the w such that dw is M dx plus N dy. Then w equals a constant is a solution. If it's not exact, find an integrating factor. Make it exact.

The trouble is that, in real life, given a particular equation, it's very, very difficult to find the integrating factor. Consequently, what one does is one categorizes various types of first order, first degree differential equations according to their structure, and shows little tricks for solving special kinds of equations-- special types where we can cut through the red tape and either find our integrating factor or a direct solution rather easily.

At any rate, that part is taken care of magnificently in the textbook. And coupled with my unique learning exercises, you will get an excellent amount of drill in how to do the mechanics. But what I wanted this lecture to do was to have you at least understand, for sure, what you meant by a solution to an equation-- what you meant by a general solution, a particular solution, a singular solution, et cetera. At any rate, then, until we meet again for second and higher order differential equations, let's just say so long for now.

Funding for the publication of this video was provided by the Gabriella and Paul Rosenbaum Foundation. Help OCW continue to provide free and open access to MIT courses by making a donation at ocw.mit.edu/donate.