

ANNOUNCER 1: The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high quality educational resources for free. To make a donation or to view additional materials from hundreds of MIT courses, visit mitopencourseware@ocw.mit.edu.

PROFESSOR: Hi. Today we're going to study something called the hyperbolic functions. And in a sense, we're going to show that in a very major way, the study of the hyperbolic functions mimics the study of the circular functions. And so in a sense, what we will do is parrot much of what we did for circular functions to wind up with our hyperbolic functions.

Since the basic difference between a hyperbola and a circle is a sign, namely 'x squared' plus 'y squared' equals 1 versus, say, 'x squared' minus 'y squared' equals 1, I've entitled today's lesson, 'What a difference a sign makes'.

Now, by way of brief review, recall that the circular functions originated as follows: we said consider the circle whose equation is 'x squared' plus 'y squared' equals 1. That's a circle over here. Then what we did was picking this as a starting point, and taking any length, 't', we wrap 't' around the circle. When the length, 't', terminated at point, 'p', we defined what? Cosine 't' to equal 'x' and sine 't' to equal 'y'. That was basically the geometric interpretation of the circular functions, after which we showed that there was a physical meaning to this as well.

Now you see what we'd like to do is the same thing only which regard now to what I would call the hyperbolic functions. And you may recall that in our lecture on circular functions, I mentioned the hyperbola very briefly. Now I'd like to go back to the hyperbola in somewhat more detail, namely we now look at the equation 'x squared' minus 'y squared' equals 1. See just to change of sign over here.

The graph of 'x squared' minus 'y squared' equals 1 is the hyperbola. And by the way, notice already one significant change when we change the sign here. Notice that in the case of 'x squared' plus 'y squared' equals 1, our curve essentially came in one piece. Now you see our curve has two different branches. And for the sake of uniformity, for the sake of well-definedness, let's pick the branch that we will deal with to be the positive branch, meaning the branch that's to the right of the y-axis.

To mimic precisely what we did in the case of the circular functions, what we do now is what?

Given any length, 't', we start at 's', mark 't' off along the upper branch of this double-valued curve if 't' is positive. Mark it off along the lower branch if 't' is negative. Wherever 't' terminates, we call that point 'p'. And then to complete our analogy with the circular functions, we define the x-coordinate of 'p' to be the hyperbolic cosine written C-O-S-H and pronounced cosh. 'Cosh t' is 'x'. And the hyperbolic sine is the y-coordinate of 'p'. And that's written S-I-N with an h, hyperbolic sine pronounced cinch. In other words, 'cosh t' is the x-coordinate of 'p'. 'Sinh t' is the y-coordinate of 'p'.

And the same way that 'x squared' plus 'y squared' equals 1 gave us the circular identity that 'sine squared t' plus 'cosine squared t' is 1, the fact that 'x squared' minus 'y squared' is 1 gives us the hyperbolic identity that 'cosh squared t' minus 'sinh squared t' is 1.

So again, notice great similarities, great differences. I'd like to make just a brief aside to point out how in one sense this difference isn't too significant. Yet in another sense, it's very significant. Notice the following: if you look at the equation, 'x squared' minus 'y squared' equals 1, and if I can draw on your previous knowledge of the square root of minus 1 just for a moment for sake of illustration. If we think of 'i' as being a number whose square is minus 1, noticed another way of writing 'x squared' minus 'y squared' is to write 'x squared' plus the number 'iy squared'. You see 'i squared' is minus 1. 'y squared' is 'y squared'. This is just another way of saying 'x squared' minus 'y squared' is 1.

Again, it's not that crucial. But because much later in the course when we study complex numbers, there is going to be a rather nice identification between hyperbolic functions and circular functions as a prelude of things to come. Notice that what we're saying is if you plot 'x squared' minus 'y squared' equals 1, if you plot that in the xy- plane, you get the hyperbola that we talked about.

On the other hand, if you change the name of the y-axis to the iy-axis, notice that with respect to this type of coordinate system, the graph would be a circle. Again, just an aside to show you structurally a connection between circles and hyperbolas in terms of complex numbers. In that way, these things out fairly similar.

Now to show you how different a sign change is, I thought you might be interested in the following little device. Take any two numbers, 'a' and 'b', no matter how different, and form two numbers, one of which is half the sum of 'a' and 'b'. And the other is half the difference of 'a' and 'b'. It's a trivial verification to show that the sum of 'x' and 'y' is 'a'. And the difference 'x - y'

is 'b'. In other words, no matter how different 'a' and 'b' are, we can always find two numbers, 'x' and 'y', such that just by changing the sign, namely if we add the pair of numbers, we get one of the numbers. If we subtract them, we get the other.

Just to make this a little bit more vivid, suppose, for example, 'a' happens to be 1,000. And suppose 'b' happens to be 4. Now there's no danger of confusing these two numbers. They're quite different. On the other hand, notice that half the sum is 502, and half the difference is 498. You see if I add these two numbers I get 1,000. If I subtract them I get 4. And therefore, notice that just by changing the sign here, I can effectively create a change from 4 to 1,000.

I just mentioned that to show why these people who say things like, all I was, was off by a sign, that making a sign mistake is very, very crucial, especially if the answer later on has to be added on to something else. But any rate, I didn't even have to give this example. Except that I thought it was a cute little aside for you to notice.

From a more practical point of view, if we just go back to these two graphs that we've drawn, notice that just changing the plus sign to a minus sign radically changed the shape of the curve that we were talking about. So at any rate then, this gives us the hyperbolic functions.

And a very natural question to raise at this time, is why the hyperbolic functions now? Why study them now? Why didn't we do it before? Why did we wait until we were dealing with exponential functions? The answer is that the exponential functions give us a very vivid way of reconstructing the hyperbolic functions. And the key lies in the following fact. It lies in the fact that the derivative of 'e' to the 'u' with respect to 'x' is 'e' to the 'u' times 'du/dx'. In particular, if I take the derivative of 'e' to the 't' with respect to 't', I get 'e' to the 't'. If I take the derivative of 'e' to the minus 't' with respect to 't', I get what? 'e' to the minus 't' times the derivative of minus 't' with respect to 't'. That's minus 1. So the derivative of 'e' to the minus 't' is minus 'e' to the minus 't'.

Well what does this have to do with anything? Notice that if I now construct a new function, 'e' to the 't' plus 'e' to the minus 't', since the derivative of a sum is the sum of derivatives, notice that this result here tells me that the derivative of 'e' to the 't' plus 'e' to the minus 't' with respect to 't' is 'e' to the 't' minus 'e' to the minus 't'.

Now let me do this operation once again. Suppose I now take 'e' to the 't' minus 'e' to the minus 't' and I differentiate that with respect to 't'. Well, the derivative 'e' to the 't' gives me 'e' to the 't' again. The derivative of 'e' to the minus 't' is minus 'e' to the minus 't'. Minus times

minus is plus. So observe that if I differentiate 'e' to the 't' minus 'e' to the minus 't', I come back to 'e' to the 't' plus 'e' to the minus 't'.

Notice already how this starts to behave, at least a little bit, like our friends sine and cosine, except again, for a little change in sign, namely if I give this function here a special name. And to make it suggestive, I'm going to call it 'c of t'. 'c' sort of to indicate cosine. But I don't want to write cosine because it's not a cosine. Let me just call this function here 'c of t' and its companion, as you may guess, I will call 's of t'.

'c of t' is 'e' to the 't' plus 'e' to the minus 't'. 's of t' is 'e' to the 't' minus 'e' to the minus 't'. This pair of functions has the property that what? The derivative of one of them is always the other, which is not quite what sine and cosine did. Remember if you differentiated the sine, you got the cosine. But if you differentiated the cosine, you didn't get the sine. You got minus the sine. So you see, again, a similarity, but not exactly the same thing.

Nonetheless we get a little bit suspicious about this. We see that the derivative of 'c' is 's', the derivative of 's' is 'c'. We remember that there was a basic identity involving sine squared plus cosine squared. We might get the premonition that maybe there's some relationship between 's squared' and 'c squared'. And one of the best ways of checking out a premonition is to just check it out. I mean let's just see what happens.

In other words, starting with this definition of 'c'-- this is a binomial-- let's square it. You see if we square it, it's what? It's 'e' to the '2t', the first term squared, plus twice the product of these two terms. But 'e' to the 't' times 'e' to the minus 't' is 'e' to the 0, which is 1, plus 'e' to the minus 't squared', which is 'e' to the minus '2t'. In other words, putting this all together, the square of 'c' is 'e' to the '2t' plus twice 1, which is 2, plus 'e' to the minus '2t'. Similarly, when we square 's', we get exactly the same thing only with a minus 2 as the middle term rather than a plus 2.

Well now we look at 'c squared' and 's squared'. And just by looking at this expression, any hope that 'c squared' plus 's squared' will give us a nice identity should vanish. Because you see, if you add these, you're going to get an 'e' to the '2t' term in here, in fact, twice 'e' to the '2t'. Here you'll get twice 'e' to the minus '2t'. And only the 2 and the minus 2 will drop out.

On the other hand, should you elect to subtract 's squared' from 'c squared', the 'e' to the '2t' terms cancel, the 'e' to the minus '2t' term cancels, and we wind up with what? 'c squared' minus 's squared' is identically 4.

And by the way, notice up to this point, no use of the word hyperbola. OK? However, if you think of the so-called unit hyperbola, the one for which 'x squared' minus 'y squared' is 1, notice that 'c squared' minus 's squared' equals 4 almost has that form. Well it has the form. All that's spoiling it is the 4 on the side. Now you could pretend the 4 was a 1. But that would be cheating. What we do instead is just divide both sides of this equation by 4, observing that 'c squared' over 4 is just 'c of t' over 2 squared. 's squared' over 4 is just 's of t' over 2 squared. In other words, dividing through by 4, this equation can be put into this form.

And now you see we satisfy the basic hyperbolic identity of this squared minus this squared equals 1. By the way, since the c's and the s's may not seem familiar to us, what we can do is what? Go back to the basic definition. 'c of t' meant 'e' to the 't' plus 'e' to the minus 't'. Therefore, half of 'c of t' is 'e' to the 't' plus 'e' to the minus 't' over 2. 's of t' was 'e' to the 't' minus 'e' to the minus 't'. Therefore 's of t' over 2 is 'e' to the 't' minus 'e' to the minus 't' over 2.

And again, the reason I bring this out, is it in most textbooks they simply define the hyperbolic functions this way. I find it a poor motivation. I don't find it easy to motivate where are the clear blue sky, you bring in a 2 in the denominator. All I want you to see in terms of motivating what's going on so far, is that if I just defined a function to be 'e of t' plus 'e' to the minus 't' and another one to be 'e of t' minus 'e' to the minus 't', I could have shown that the first squared minus the second squared was 4. And the motivation of putting the 2 in the denominator is for no greater reason than to help us satisfy the equation 'x squared' minus 'y squared' equals 1.

In fact, this is now the connection between what we were doing a few months ago and what we're doing now. Namely if I now call this 'cosh t' and I call this 'sinh t', do I get my basic hyperbola? Namely, suppose I take a curve whose equation is given parametrically by 'x' equal 'cosh t' and 'y' equal 'sinh t', where cosh and sinh are now defined as over here without any reference to the hyperbola, in the same way that we could define the circular functions in terms of calculus without having to appeal to the circle.

What I'm saying is, suppose we had never mentioned the word hyperbolic function up to now. We've developed a function called 'cosh t', a function called 'sinh t' just from these two relationships here. If I eliminate the parameter 't' here, namely if I square both sides of each equation and subtract, I get 'x squared' minus 'y squared' is 'cosh squared t' minus 'sinh squared t'. But that I already say is 1. In other words, this parametrically gives us the hyperbola 'x squared' minus 'y squared' equals 1.

By the way, let's make another observation. Notice that the exponential could never be negative. See e^{-t} is $1/e^t$, which is still positive. Notice that the cosh can never be negative. See, e^t is a positive number. e^{-t} is a positive number. Therefore, notice that in this particular relationship, not only is the curve given by $x^2 - y^2 = 1$, but since the cosh can never be negative, it's also characterized by x is positive. And you see that tells us what? That what we're getting this way is not the entire hyperbola, but just the branch for which x is positive, which is exactly what we were talking about earlier in the lesson.

Again, I leave further details of this to the exercises. All I want to do in today's lesson is to get the flavor of what's going on, how we can invent the hyperbolic functions without reference to a hyperbola until we're all done. See, I hope you see this juxtaposition. We've done two things here. One is we've started with the hyperbola and showed how we could invent this construction. The other is we started with the exponential and showed how we could construct two functions of the exponential which had the property that they could be identified with the thing that was called the hyperbolic cosine and hyperbolic sine. In other words, we could take either of the two approaches and derive the other. But I don't want to put in all of those details here. OK?

At any rate, to show you where calculus comes in, in terms of what I meant by saying that as far as this course is concerned, we have now finished the groundwork both with respect to differential and integral calculus. Suppose, for example, I now want to define the hyperbolic tangent. See y equals hyperbolic tangent x . How could I do that? Well, I mimic the circular function definition. I say, OK, I will define the hyperbolic tangent to be the hyperbolic sine over the hyperbolic cosine.

By the way, in the same way that one can show in terms of geometric constructions on the unit circle how one constructs the tangent and the cotangent and the cosecant and the secant, et cetera, once sine and cosine are given. It can also be shown that we can construct these things in terms of geometry from the hyperbola. But because that gets computationally involved, I've also saved those aspects for the exercises. I'm not going to worry about how you construct the hyperbolic tangent given the hyperbolic sine and the hyperbolic cosine right now. All I'm saying is suppose you've made up this definition. And now if somebody says gee, I would like to find the derivative of the hyperbolic tangent of x .

What I want to point out is that the recipes that we need are no different from anything we've had before. Namely, since the hyperbolic tangent of 'x' is 'sinh x' over 'cosh x', I use the same quotient rule as I used before. Namely, to differentiate the hyperbolic tangent, I differentiate this quotient. How do we differentiate a quotient? It's the denominator times the derivative of the numerator. Well the numerator is 'sinh x'. The derivative 'sinh x' is 'cosh x'. That gives me a 'cosh squared x' over here minus the numerator, which is 'sinh x' times the derivative of the denominator. Well the derivative of 'cosh x' is 'sinh x'-- that give me another factor-- over the square of the denominator.

Now my basic identity is that cosh squared minus sinh squared is 1. So this is just 1 over 'cosh squared x'. And if I want to write this all on one line, I could invent the notation that the hyperbolic secant is 1 over the hyperbolic cosine. In other words, I could invent a notation like this. But again, the important point is not so much what this answer is, but how I can derive it from first principles.

And by the way, in terms of the inverse derivative, once I know that the derivative of hyperbolic tangent is a square of the hyperbolic secant, I can immediately add to my list of inverse derivatives. Namely I just have to write it down. See there's an integral formula that I now have. But again, all I want to emphasize here is how the details take care of themselves in the same way as always once we have our basic definitions.

It might be interesting just in terms of cementing down what sinh and cosh look like, in the same way that we have talked about sine and cosine, it might be nice to graph sinh and cosh. And again notice, in terms of curve plotting, how do we plot 'y' equals 'f of x'? The general procedure was what? Always take the first and second derivatives so that you can see what the slope is, what the concavity is, et cetera. The derivative of cosh is sinh. The derivative of sinh is cosh. By the way, again, if these terms seem alien to you, you can always rewrite them in terms of the basic definition, in terms of 'e' to the 'x' and 'e' to the minus 'x', and carry out the differentiation in a straightforward way.

Well here's the interesting point. Let's plot 'y' equals 'cosh x'. Notice first of all that 'cosh x' is an even function. If I replace 'x' by minus 'x', all I get is what? 'e' to the minus 'x' plus 'e' to the 'x' over 2, which is the same thing as I have over here. It's going to be a symmetric function with a curve with respect to the y-axis.

But again, what do I do? I compute the derivative. I can find out where the derivative is 0. It's 0

when 'e' to the 'x' equals 'e' to the minus 'x'. That happens only when 'x' is 0. When 'x' is 0, this is 'e' to the 0 plus 'e' to the 0. That's 1 plus 1, which is 2 over 2, which is one. So in other words, the derivative is 0 when 'x' is 0 and 'y' is 1.

The second derivative is cosh. We've already seen that this can never be negative. So the second derivative is always positive. That means that the curve is always holding water. Putting all this together, the curve 'y' equals 'cosh x' looks something like this.

It's a dangerous thing to say it sort of resembles a parabola. It's nothing like a parabola except what I mean is it has this type of shape. Notice it does not oscillate. It is not an oscillating function like the cosine. It doesn't act periodically. This thing just keeps going like this. All right?

On the other hand, how can we plot 'y' equals 'sinh x'? Well, I guess we come right back to here and work from here. The derivative of sinh is cosh. But cosh is always at least as big as 1 we found out. Therefore, the slope of 'sinh x' is always at least as big as 1. And it's equal to 1 when 'x' is 0.

Putting all these details together, what we find is what? First of all, the sinh is an odd function, meaning that if we replace 'x' by minus 'x', we change the sign. So this is 'e to the minus x' minus 'e to the x', which is a negative of what we have over here. But those, again, are details which are easy for you to fill in. The graph 'y' equals 'sinh x' looks something like this. In other words, it's a very steep curve. It's spilling water and rising here, holding water and rising here. The curve is always rising.

Again, another interesting thing to observe here is that notice that when 'x' is very, very large, 'e' to the minus 'x' becomes negligible. And if 'e' to the minus 'x' is negligible, notice that both sinh and cosh behave like $\frac{1}{2}$ 'e to the x'. In other words, this term tends to drop out.

And just to show you the contrast here, I've sketched in the curve 'y' equals ' $\frac{1}{2}$ 'e to the x"' to show you how it splits the difference for large values of 'x' between these two in a way. And as 'x' gets larger, both of these curves converge on 'y' equals ' $\frac{1}{2}$ 'e to the x". Again, I simply want to mention enough here so that you get the idea of how we can apply the same old calculus to our new function.

What I'd like to do in closing is to try again from our point of view of being engineeringly oriented to show why the hyperbolic functions might have occurred in nature, in the same way

that we tried to show that the circular functions had a natural outgrowth in terms of nature.

You recall that the circular functions were motivated in terms of simple harmonic motion. The circular functions would have been invented when we talked about motion, in which the acceleration was proportional to the displacement but in the opposite direction. Let's see what happens when we work with the hyperbolic functions.

Suppose now I have a particle moving along the x-axis. OK? And I know this position at any time, 't', is \sinh of 'kt' where 'k' happens to be a constant. What is dx/dt ? The derivative of ' $\sinh u$ ' with respect to 'u' is ' $\cosh u$ '. So the derivative of ' $\sinh kt$ ' with respect to 't' would be ' $\cosh kt$ ' times the derivative of 'kt' with respect to 't', which is just 'k'. In other words, the chain rule again, OK?

Now knowing that dx/dt is ' $k \cosh kt$ ', " $d^2 x / dt^2$ " would be what? The derivative of ' $\cosh kt$ ' is ' $\sinh kt$ ' times the derivative of 'kt' with respect to 't'. That gives me another factor of 'k'. And therefore the second derivative of 'x' with respect to 't' is " $k^2 \sinh kt$ ".

Now remember what ' $\sinh kt$ ' is. We started with the knowledge that ' $\sinh kt$ ' is 'x'. Therefore replacing ' $\sinh kt$ ' by what's it's equal to, we wind up with what? The second derivative of 'x' with respect to 't' is equal to ' $k^2 x$ '. See, no minus sign here.

What happens when we start with a \sinh ? If an object we moved in according to the rule, 'x' equals ' $\sinh kt$ ', its acceleration would be proportional to the displacement but in the same direction. You see what that means? That means the bigger the displacement, the bigger the acceleration so the faster the object keeps moving.

In other words, this thing would move at a faster and faster right away from the origin, et cetera. That's not our main concern here. What is interesting though from a physical point of view, and we can use this as a summary from a physical point of view, is that the hyperbolic functions serve as a solution to the equation " $d^2 x / dt^2$ " equals ' $k^2 x$ ' where 'k' is a constant.

The circular functions, on the other hand, solve the equation " $d^2 x / dt^2$ " is minus ' $k^2 x$ '. In other words, from this point of view, notice again, from a physical point of view that the difference between the circular functions and the hyperbolic functions are again just the difference in sign.

As I say, this will be emphasized more in the exercises. The obvious next lesson, I think you

can see it coming by this stage in the game, will, of course, be the inverse hyperbolic functions. And we'll see why those are important next time. But until then, good bye.

ANNOUNCER 2: 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.