

ANNOUNCER: 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 MIT OpenCourseWare at ocw.mit.edu.

HERBERT

GROSS:

Hi, our topic today sounds a little bit naughty at first glance. It's called improper integrals. And rather than give a long speech about why they're called improper, let me try and start our lecture a little bit different from our usual approach, by giving you a little bit of a mathematical calculus riddle. Namely, let me just say this. We'll call our lecture today improper integrals, and we'll start off by saying find the flaw.

I would like to integrate 'dx' over 'x squared' from minus 1 to 1. Using the first fundamental theorem of integral calculus, I say this is 'G of 1' minus 'G of minus 1' where 'G prime' is any function whose derivative is '1 over 'x squared''.

Well, in particular, do I know a function whose derivative is '1 over 'x squared''? Well, I guess minus '1 over x'. In other words, the derivative of 'x' to the minus 1 is minus 'x' to the minus 2. With the minus in front that makes it plus. 'x' to the minus 2. That's plus '1 over 'x squared''. In any event, I can then compute that 'G of 1' is minus 1. 'G of minus 1' is 1. Therefore, 'G of 1' minus 'G of minus 1' is minus 1 minus 1, is minus 2. Therefore, the integral from minus 1 to 1, 'dx' over 'x squared', is minus 2. And the question is, find the flaw. Find the flaw. And since again, we're on limited time, let me, if you haven't already discovered the flaw, at least point out why you should be suspicious of a flaw.

Remember, we visualized this as an area under a curve. In other words, in a way, the base of our region is the closed interval from minus 1 to 1, and the height of our region at any point is given by '1 over 'x squared''. Well, look at '1 over 'x squared''. Certainly, '1 over 'x squared'' can't be negative because for any real number 'x', its square is non-negative. And one over a non-negative number is certainly non-negative. In particular then, since my region never goes below the x-axis, I would expect that the area represented by this should be positive. In other words, this is the sum of positive numbers taken in the limit. In other words, the fact that the integrand is non-negative tells me that whenever the integral turns out to be, it must be at least as big as 0, and minus 2 does not fulfill that criteria.

In other words, something must be wrong here because we get a negative answer in a situation where a negative answer is preposterous. And the reason that I wanted to pick this approach is I think that the best way to motivate that something is wrong is precisely by winding up with a preposterous answer in a situation where we know the answer is preposterous. Then we know we must be on our guard to see what went wrong.

And what did go wrong? Well, the key point is this. And again, it's something that we've said many, many times. But I think it takes a counter example before it really sinks in. People are always quoting the first fundamental theorem in general. It's easy to remember that the integral from 'a' to 'b', " $\int_a^b f(x) dx$ " is ' $G(b)$ minus ' $G(a)$ ', where ' G' ' equals ' f '. But the thing that we must be extremely careful of is that this definition requires that ' f ' be at least piecewise continuous on the interval from 'a' to 'b'.

Remember when we took the limit and put the squeeze on, we required that there be continuity, so we could pass to the limit. The important point is in this particular example that we're dealing with, observe that " $1/x^2$ " is infinite. I write this in quotation marks to point out you have a $1/0$ form, which we usually abbreviate as infinity. In other words, " $1/x^2$ " increases without bound as ' x ' approaches 0. And notice that 0 was within the limits of our integration from minus 1 to 1.

At any rate, without further ado, this leads to a fundamental definition. Which in terms of today's lesson, I call definition number 1. And that is the integral from 'a' to 'b', ' $\int_a^b f(x) dx$ ' is called 'improper'. And there are going to be two types of improper integrals as we'll find out very shortly. But it's called improper of the first kind, if and only if ' f ' is infinite for at least one ' c ' in the interval from 'a' to 'b'. And again, let me caution you about this and I'll summarize at the end of the lecture again to remind you.

If you look at this particular integral, it doesn't look at all suspicious. In fact, that's where the trap came in in this problem over here. The reason that we wound up with minus 2 is that we just integrated " $1/x^2$ ", forgetting that there was a point at which the integrand was infinite in the given interval.

And so the question that comes up is, how shall we handle this if we have an infinity in here? And the answer is, let's work around the point at which the integral or the integrand is infinite. Let's see what that means. And this again, is a continuation of the definition, I guess.

If ' c ' is the only point in the closed interval from 'a' to 'b' at which ' f ' is infinite, we defined the

integral from 'a' to 'b', $\int_a^b f(x) dx$ to be the following limit. It's the limit as 'h' approaches 0 through positive values. The definite integral from 'a' to 'c' minus 'h'. In other words, you're stopping somewhere before you get to this bad point, evaluating $\int_a^{c-h} f(x) dx$. And there's no harm done in here because you see, $f(x)$ is infinite only at $x = c$ in this problem. Consequently, if you stop prior to 'c', $f(x)$ is at least piecewise continuous every place in here. Then you jump over the bad point, pick up again at 'c plus h' and say plus the integral from 'c plus h' to 'b', $\int_{c+h}^b f(x) dx$. This limit as 'h' approaches 0 through positive values.

Now again, this may sound very abstract. So to give this a more realistic meaning to you, let's interpret this again in terms of an area. What we're saying is this pictorially. We have a function which is piecewise continuous at least on the entire interval from 'a' to 'b', except at the point 'c' where we have an infinite discontinuity, an infinite jump. What we do is, we pick a positive number 'h' and knock off the interval 'c minus h' to 'c plus h' surrounding 'c'. We then compute the area. We then compute the area of this shaded region. That's exactly what's inside the brackets over here. You see, notice that this sum is what? It's the sum of two integrals, one of which names this area, and the other of which names this area. And what we say is, let's put the squeeze on and see what limiting value this area approaches as we squeeze all the space out around 'c'. In other words, what's really happening over here?

What we're really saying is, why can't we treat an infinite region the same as if it were a finite region taking the limit as the size of the strip in here goes to 0? And here's the key point. You see, as 'h' approaches 'c' from both sides, notice that the width of this region approaches 0. By the same token, the height of the region is approaching infinity. In essence, what we're running into again, is the problem of determining the area of an infinite region involves the study of infinity times 0.

In fact, the question sort of is, does 'h' go to 0 faster than ' $f(c \pm h)$ ' goes to infinity? In fact, I think that's an important enough observation, so I've written it down here to make sure that you notice it. The question centers about the question, the problem of whether 'h' approaches 0 faster than-- whatever this means. Faster than ' $f(c \pm h)$ ' approaches infinity.

In our particular example, this didn't happen. Now again, if you're afraid of the symbol infinity, let's deal with the finite case and take the limit later. In other words, instead of looking at the integral from 0 to 1, $\int_0^1 dx/x^2$, let's look at the integral from 'h' to 1. Integral from 'h' to 1, $\int_h^1 dx/x^2$, and then later we'll take the limit as 'h' approaches 0.

In particular, if this is my region 'R', namely bounded above by the curve 'y equals 1 over x squared', below by the x-axis. On the left by the line 'x equals h' and on the right by the line 'x equals 1, the area of the region 'R' is written what? Integral 'h to 1, 'dx' over 'x squared'. Notice that as long as 'h' is not 0, our integrand is continuous. Therefore, we can use the first fundamental theorem to evaluate it.

It's what? It's minus '1 over x' evaluated between 'h' and 1. Putting in the upper limit gives me a minus 1. Putting in the lower limit gives me minus '1/h'. But I'm subtracting that, so it becomes plus '1/h'. In other words, the area of my region 'R' is $1/h$ minus 1.

Notice then that 'A sub R' as 'h' approaches 0 from the right becomes the limit of '1/h' as 'h' approaches 0 from the right minus 1. And the limit of '1/h' as 'h' approaches 0 from the right is clearly infinity. In other words, as 'h' gets arbitrarily close to 0 through positive values, '1/h' increases without any bound. In other words, what this means in plain English is that we can make the area of the region 'R' as arbitrarily as large as we wish just by choosing 'h' to be sufficiently close to 0 here. You see, in other words, as 'h' approaches 0, what's happening here is that this curve is not approaching the y-axis fast enough to allow us to get a finite area. And by the way, let me make one little aside comparing this with our previous answer.

You may recall, and let me just pull this board down here to remind you of that. You may recall that when we worked this problem before, we got the wrong answer, negative 2. OK, now look it. In terms of my problem over here, we went from minus 1 to 1 so this is only 1/2 of the region that we're talking about. Notice that the minus 2 comes from this. See twice minus 1 by symmetry.

The interesting thing is notice here that when you take the limit-- well, forgetting that there's an infinite discontinuity here, when you let 'h' approach 0, you tend to say let's forget about the function of 'h' in here. Notice here as 'h' approaches 0, this thing went to infinity. By ignoring it, that's where you got the answer minus 2. In other words, by assuming that as 'h' got small, the function of 'h' got small, we threw out an infinite part of our answer. In other words, this was the thing that we had to worry about. In this particular case, we slurred over something that wasn't going to 0 fast enough.

By the way, let me caution you and point out that in this particular example it happened that infinity times 0 became infinite. It is possible, however, that as 'h' approaches 0, it may happen

that 'h' approaches 0 fast enough, so that we do get a finite area.

To show you this, let's just change the problem slightly. Instead of taking 'y' equals '1 over 'x squared'', let's take 'y' equals '1 over the 'square root of x''. The same problem, and by the way, here's the subtlety of this again.

If you were to sketch 'y' equals '1 over the 'square root of x'' alongside of 'y' equals '1 over 'x squared''-- in fact, fast enough so that even though we get an infinite region, it has only a finite area. In particular, here's what I'm saying now. Let 'R' be the region bounded above by 'y' equals '1 over the 'square root of x'', below by the x-axis. On the left by 'x' equals 'h' and on the right by 'x' equals 1.

Again, since the only place that '1 over the 'square root of x'' has an infinite discontinuity is when 'x' is 0, and since 'h' is greater than 0 here, we have a continuous integrand. The area of the region 'R' therefore is given by what? Any function whose derivative is '1 over the 'square root of x'' evaluated between 'h' and 1. Again, sparing you the details. Observe that if you differentiate twice 'x to the 1/2', the 1/2 and the 2 cancel and you get '1 over the 'square root of x''. In other words, 2 square root of 'x' is the function whose derivative is '1 over the 'square root of x''.

If I now evaluate that between 1 and 'h', I get 2 minus 2 'square root of h'.

If I now take the limit of 'A sub R' as 'h' approaches 0, notice that what happens here is that as 'h' approaches 0, so does the square root of 'h'. And the area turns out to be 2 in the limit. In other words, the area of the region 'R' in this particular case, is also a function of 'h'. But the part that depends on 'h' goes to 0 as 'h' goes to 0. In fact, notice in this particular case, and why I couldn't have started with this example is that if I took this example and you did it the incorrect way, in other words, if you failed to notice that '1 over the 'square root of x'' was discontinuous at 'x' equals 0, and you just did this problem mechanically, and evaluated this between 0 and 1, you would have got the right answer. And the reason that you would have got the right answer, is that in the limit process, notice here that the part of the answer that depends on 'h' does actually go to 0 in the limit.

At any rate, using all of this as motivation, we come to a new piece of terminology that we should define. Namely, if 'f' is infinite at 'x' equals 'c', where 'c' is someplace in the interval from 'a' to 'b', then the improper integral 'a' to 'b', 'f of x' dx' is called convergent if the particular limit that we talked about previously in the definition of how you compute integral from 'a' to 'b',

" $\int_a^b f(x) dx$ ", if that particular limit exists, meaning it's a finite number, then we call the integral a convergent improper integral.

On the other hand, meaning what? If this limit doesn't exist, if for example, the limit is infinite, then we call the integral divergent. In other words, in our first example, the integral was divergent. In the second example, the integral was convergent.

By the way, it's rather interesting to point out too-- I shouldn't say it's rather interesting. It's just my excuse of I don't know of any other way of motivating improper integrals of the second kind. But the thing that's worth observing is this. When we evaluated the area under the curve here from 0 to 1, we elected to pick strips like this. Remember this is why we got into that infinite discontinuity; we picked out strips to look like this.

The question that comes up is, why couldn't we have picked our strips to be horizontal? In other words, if we had in terms of say, inverse functions, inverted the role of 'x' and 'y', why couldn't we have computed our area by picking strips like this and adding them all up from 1 to 'b', and then taking the limit as 'b' went to infinity? In other words, if I wanted to compute the area of this region using horizontal strips, I would have had what? I would have had the area of the region 'R' is the integral from 1 to 'b'. And remember, if 'y' equals '1 over x squared', 'x' is '1 over the square root of y'. Therefore, the area of the region 'R' is the interval from 1 to 'b', 'dy' over the square root of 'y'. Notice that the square root of 'y' is certainly continuous. '1 over the square root of y' is continuous in this interval. At any rate, I could then have evaluated this integrand here and taken the limit as 'b' approached infinity.

And by the way, the abbreviation for writing the limit of the integral 1 to 'b' as 'b' approaches infinity, is to simply write integral from 1 to infinity 'dy' over the square root of 'y'. And this is all this thing means. For example, in this particular case, notice that as 'b' goes to infinity, so does the square root of 'b'. Therefore, $2\sqrt{b} - 2$ goes to infinity as 'b' approaches infinity. And we find by a second method that the area of our region is infinite. And again, we have another kind of improper integral.

But structurally, this improper integral is completely different from the improper integral that we had before. Namely, before we had an improper integral where the integrand became infinite, even though the limits of integration stayed finite. Notice what happens in this case.

In this particular case, notice that the integrand stays finite. It's the limits of integration that become infinite. In other words, this leads to our next definition of a new type of improper

integral. Namely, the integral say, from 'a' to infinity "f of x' dx'. Whereby this we mean the abbreviation limit as 'b' approaches infinity, integral from 'a' to 'b', "f of x' dx', where 'f' is continuous in this entire range. In other words, for 'x' greater than or equal to 'a'. That type of an integral is called an improper integral of the second kind. In other words, in the first kind, the integrand becomes infinite, but the limits of integration are finite. In the second kind, the integrand stays finite, but the limits of integration take on an infinite range. And the important point is that these two are very closely connected. In other words, in terms of the example that I just showed you, I wanted you to see how you can get from an integral of the first kind to an improper integral of the second kind just by replacing your vertical strips by horizontal strips.

In fact, in terms of my first example over here notice that indeed, '1 over 'x squared'" and '1 over the 'square root of x"', where 'x' is at least as big as 0, are inverses. By the way, the reason I say 'x' is, in fact, bigger than 0, otherwise I'll have a 0 denominator here. The reason I point this out is that if you don't presuppose that x is positive, remember, the square root function or the squaring function is not 1:1, so I have to be very careful when I talk about inverse functions. The thing I want to see however is, the graph of 'y' equals '1 over 'x squared'" is the inverse of the graph 'y' equals '1 over the 'square root of x"'. In other words, essentially, by interchanging the role of the x- and y-axes, you go from one graph to the other.

You see, in particular, the integral 1 to infinity, 'dx' over 'x squared'. Namely, finding the area of this particular region is the inverse of finding the area of this particular region. In other words, notice that 'y' equals '1 over 'x squared'" behaves for large values of 'x' the same as 'y' equals '1 over the 'square root of x" behaves for small parts of the values of 'x'. Again, let me point out that much of this will be done much more hard core in our learning exercises and in the reading material and the like. And so I pass over some of this in a fairly quick way. But what I want to make sure is clear before we begin the exercises and the reading material is that we have a good idea as to what improper integrals mean.

At any rate, continuing on this way, what do I mean by the integral from 1 to infinity, 'dx' over 'x squared'? Well, it's an improper integral of a second type because the integrand is finite between 1 and infinity. The limits of integration span an infinite space over here. At any rate, it's what? The limit as 'b' approaches infinity, integral from 1 to 'b', 'dx' over 'x squared'.

At any rate, what is the integral of '1 over 'x squared"'. It's minus '1/x'.

If I evaluate that that's minus "1 over b' plus 1'. And if I now take the limit as 'b' approaches

infinity that answer is 1. In other words, what this means pictorially is the following. If I look at the area of this region 'R' for any value of 'b' whatsoever, that area is equal to 1 minus '1/b'. In other words, no matter how big 'b' becomes, the area of this region 'R' can never be as big as one unit. And in the limit, as 'b' goes out as far as we wish, what we're really saying is what? That the area of this region 'R' can be made arbitrarily close to 1 just by choosing 'b' to be arbitrarily large.

And by the way, this you see leads to an interesting observation. Namely, at this end of the spectrum, the curve was not approaching the y-axis fast enough to make this integral finite. Remember, the integral from 0 to 1, 'dx' over 'x squared' was infinite. However, for large values of 'x', the curve does approach the x-axis fast enough, so that not only is the enclosed area finite, it can't even be bigger than 1. I thought that might be an interesting example for you to think about for a while.

Another type of interesting example worth thinking about for a while is this second example. Let's take the region 'R' to be bounded above by 'y' equals '1/x', on the left by 'x' equals 1, on the right by 'x' equals 'b', and below by the x-axis. And let's compute the area of the region 'R'.

Again, the area of the region 'R' is a definite integral from 1 to 'b', 'dx' over 'x'. And that's equal to the natural log of 'b'. Namely, this integral is log absolute value of 'x'. Log absolute value of 'b' is just 'log b'. Log 1 is 0. So the area of the region 'R' is 'log b'. As 'b' increases without bound, its log increases without bound, we find that the area of the region 'R' is infinite. In fact, another way of saying this is what? This is the integral 1 to infinity 'dx' over 'x'. And this is, therefore, what? A divergent improper integral of the second type. All right, so far so good at any rate. Let me review our volumes of revolution and rotation.

Let's take the region 'R' and rotate it about the x-axis. If we rotate 'R' about the x-axis, remember a volume of revolution is just pi 'y squared' 'dx' from 1 to 'b'. Well, my 'y' is '1/x'. That's '1 over x squared'. The integral of '1 over x squared' is minus '1/x'. Notice then as before, the volume when I rotate 'R' about the x-axis is just pi times '1 minus '1/b''. In other words, what this says is as 'b' gets very, very large, so does 'R'. Yet, if I rotate 'R' about the x-axis, no matter how big 'b' is the volume is just pi times '1 minus '1/b''. As 'b' approaches infinity, this goes to 0 and my volume is just pi. In other words, this area gets infinitely large. Yet, the volume generated by rotating that area remains finite. Again, another idiosyncrasy of working with improper integrals and infinite regions. You've got to be very, very careful about how rapidly something approaches 0 or infinity because infinity times 0 is indeterminate. In

other words, here is a very important observation, an infinite area can enclose or can generate a finite volume. Be very, very careful in dealing with improper integrals.

And by the way, let me make one computational aside. In trying to determine whether a particular improper integral is convergent or not, we do not have to be able to compute the limit itself. I mention this because it's very important. Sometime you'll be given an integral which you don't know how to integrate directly. A typical example that usually comes up, integral "e to the minus 'x squared' dx", say between 1 and infinity. And you want to know whether this converges or not. You don't care what the limit is because once you know it converges, you can approximate it to as great a degree of accuracy as you wish. So the question is, does this converge or doesn't it? And again, the answer is, let's see what happens on this particular range, for example. In other words, we know, for example, that if 'x' is greater than or equal to 1 for numbers in magnitude at least as big as 1, the square exceeds the number.

In particular by the way, if 'x squared' exceeds 'x', minus 'x squared' cannot exceed minus 'x'. In other words, on this particular interval that we're dealing with, say from 1 to 'b', as 'b' approaches infinity, 'e to the minus 'x squared'" is less than minus 'x'. Therefore, because the exponential function is 1:1, "e to the minus 'x squared' dx' from 1 to 'b' cannot exceed the integral from 1 to 'b', "e to the minus x' dx'. In other words, since this exponent is less than this exponent, this integral is less than this one. However, it turns out that we know how to evaluate this one. How do we integrate 'e to the minus x'? It's just minus 'e to the minus x'. And if we evaluate that between 1 and 'b', we find that we get '1/e' minus '1 over 'e to the b' power. 'b' is some positive number. This thing here is less than infinity.

In fact, as 'b' approaches infinity, this approaches 0, and the limiting value is '1/e'. In other words, what we now know is, is that whatever this improper integral is, it cannot exceed '1/e' even though we don't know what the exact answer is. Pictorially, you see what happens over here is that the curve which I've drawn in the accentuated chalk, this curve is 'y' equals 'e to the minus 'x squared"'. The light curve is 'y' equals 'e to the minus x'. Notice that these two curves crisscross when 'x' equals 1. And all we're saying is that since the white curve chops off a finite area and the accentuated curve, the dark curve is inside the white one, certainly the area enclosed by the black curve cannot exceed that of the white curve. In other words, if this region here approaches a finite value as 'x' goes to infinity, so must the other region. And this is essentially the theory behind improper integrals.

In closing, there is one thing I would like to say about this. It's a short summary, but here's the whole idea. It may be difficult to evaluate improper integrals of a second kind. But in the certain manner of speaking, they're not dangerous.

In what manner of speaking? Well, when you see the limits on the integral sign like a to infinity, or minus infinity to b , or minus infinity to infinity, you're on the lookout. You're on the lookout to be careful about something going wrong. There's something about the symbol infinity, which hopefully strikes a certain fear or respect in your minds when you work mathematically. See, we have a warning to be ware in this case. However, the dangerous part is that when you're given an improper integral of the first type, and I don't know the best way of saying this other than to say perhaps, that mathematics has no built-in burglar alarms. What I mean by that is if you're supposed to multiply 5 and 3 to get the right answer to a problem, and instead you add 5 and 3, you're going to get 8 for the answer even though it's incorrect. There's not going to be a buzzer that goes off and says, please do not add these two numbers.

In other words, if you go ahead over here and you forget that f of x may be infinite and you go ahead and integrate this using the first fundamental theorem, you're going to get an answer just like we did in our beginning exercise at the beginning of the lecture. It happened to be the wrong answer. So in other words, the word of warning here is whenever you see an integral with finite limits of integration, check the integrand and make sure it doesn't blow up, it doesn't become infinite.

At any rate, with these words we finish our block on integration. And until next time, good bye.

ANNOUNCER: 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.