

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 view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at ocw.mit.edu.

PROFESSOR: Hi. Today we, in a manner of speaking, start the second part of this part of our course that what we have done up to now is hopefully to have presented the concept of partial derivatives with all the trimmings.

And now what we would like to do is devote the remainder of the course to what we could call selective topics, in which the partial derivative plays a fundamental underlying thread. In other words, a situation in which we will now pick specific areas of calculus, and show how what we've learned about partial derivatives can be applied to these particular areas.

Now, the first natural area, which would suggest itself, it seems to me, would be the concept of integrating a function of several variables. Since after we study derivatives, we next associated the definite integral with an antiderivative, it might seem that a very natural inquiry now is to discuss the concept of infinite multiple sums as being the analog of the definite integral in our study of calculus of a single variable.

Now, because these topics that we're picking have great application and have been studied in great depth, we by necessity will sort of skim the surface of many of these topics, so that we can present as many of the key points as we can in the time allotted for the course.

We will make the lectures hopefully as short as possible, as pungent as possible, and leave most of the details to the learning exercises, and to the supplementary notes. At any rate, what our lecture today is called is double sums, or perhaps a better word would be multiple sums.

But we will concentrate on the case of double sums for the usual reason that

namely, in the case of two independent variables, we have convenient pictures still left at our convenience. Now the idea is, let's consider a region, R . Say the square, the unit square on the coordinate axes with vertices $0, 0, 1, 0, 1, 1,$ and $0, 1$.

Now, what we're going to do now is imagine that R is a thin plate. And what we would like to do is to find the mass of this plate R , assuming that the plate has a variable density. In other words, what we want to do is find the mass of the plate, R , if its density row at the point x, y and R is given by row of xy equals x squared plus y squared.

You see, by the way, the reason I picked this particular density of distribution is you may notice that x squared plus y squared quite simply is the square of the distance of the point x, y from the origin. And this will make it very easy later for me to find where lowest densities occur, where greatest densities occur. Whereas this may not seem like a very applied problem.

Notice that if you want the geometric picture of what's going on here, the idea here is that the density is varying regularly as we move away from the origin. In other words, the density is the square of the distance of the points from the origin, which means that the density is constant on any circle centered at the origin.

Again, the key building block that we're going to use is again, analogous to what we used in calculus of a single variable. For example, in calculus of a single variable, when we said that distance was equal to rate times time, it was assumed that the rate was constant. In a similar way, notice that if we assumed that the density were constant.

If the density were constant, then the mass would just be the density times the area. Again, most of us are used to thinking of density in terms of three dimensions, that the mass is the density times the volume. I prefer to pick a thin plate here, simply to utilize a simple diagram. But I did not want to get involved with three dimensional diagrams.

Another drawback, perhaps, to the system is that you may be wondering why we're

not using the counterpart of what area was in calculus of a single variable. Why aren't we talking about the volume in two dimensions? Why couldn't we somehow visualize the volume problem here? Again, if that's what you'd like to do, notice that we could graph the density as a function of the point x, y in the region R .

Plot the density in the direction of the positive z -axis, in which case the problem that we would then be trying to solve would be to find the volume of a solid that you get by taking the parallelepiped, whose base is the region R , and intersect that with the surface z equals x squared plus y squared. Again, we will leave these details for the exercises.

But the surface z equals x squared plus y squared as you look in along the z -axis here. There's a parabolic bowl opening up. In other words, z equals x squared plus y squared. If you take z equals a positive constant, in other words slice this parallel to the xy plane. If z is a positive constant, x squared plus y squared equals a positive constant is a circle.

In any rate, just to summarize that, the geometric equivalent to our problem is find the volume of a solid s if s is the parallelepiped with the base R where R is just as given, and the top meaning it's intersected with the surface z equals x squared plus y squared. And again, let me mention as an aside that the same analogy appears in calculus of a single variable.

Namely, if I write down the definite integral from 0 to 1 x squared dx , I may view this, well, in many ways, but two of the major ways in which I may view this, the major one that we emphasized in our course was as an area. I can also view this as a mass. More specifically, let me take the region R as shown here.

Then you see one interpretation of the integral from 0 to 1 x squared dx is that it's the area of the region R . Another interpretation is imagine that you have a very thin rod geometrically of uniform shape. In other words, it's essentially a segment of the x -axis from 0 to 1. And suppose that the density of this rod is a function of position x , say row of x is x squared. Then notice in a manner analogous to what we were just talking about, the mass of this rod would be found by integrating x squared from

0 to 1.

By the way, notice again a very important thing that comes up here, and that is that these two diagrams may be used for the same problem. You see, in many cases when one has a density distribution like this, one visualizes the density being plotted. In other words, the density row of x equals x squared is plotted as the curve y equals row of x , in this case, y equals x squared.

And we then think of a way of being able to visualize what the density is doing here in terms of the shape of the curve here. That's exactly what we were doing earlier in the lecture, when we said that we can visualize the density row of xy as a function z equals row of xy . The analogy being that with two independent variables, we wouldn't have a curve, but rather we would have a surface.

By the way, again, notice how this emphasizes the fact that the definite integral is in reality an infinite sum. That it is convenient in many cases to view this limit of a sum in terms of an area under a curve. But that the definite integral itself exists for more problems than just for computing areas.

Now at any rate, all we're trying to point out here is that the same analogy will exist when we go to functions are several variables. If in fact we now return to the problem stated at the beginning of the lecture, we take our region R . Remember what this is, now. It's these square with vertices at these points. We know that the density is given at any point x, y by x squared plus y squared.

And what we'd like to do now is find the mass of this particular plate. Now, just as in calculus of a single variable, the argument that we use now is nothing more than if we took a small enough region here, we could assume that the variable density was essentially constant. In other words, this is how one makes physical approximations. We break this thing up into small pieces.

Somehow or other find the mass of each small piece, and add those all up. The key point being if the pieces are small enough for an approximation, we can say that the density is essentially constant. By the way, in the same way that we had to learn the

sigma notation back in part one of our course, we will now have to somehow or other learn to master the notation for double sums.

The idea again being this, if I partition this into n equal parts, and the segment into m equal parts, and if I call a typical x segment, Δx_i , and a typical y segment, Δy_j , then to match up in coordinate fashion what the mass of this little piece would be, I would call that the ij mass corresponding to the i partition, and the j partition. And the idea is I'm going to add all of these up.

That would motivate the notation called the double sum. Namely the way you read this is you say, let me first pick a fixed value of j , and for that fixed value of j , I will sum up all these pieces as i goes from 1 to n . What that means pictorially is this. I will pick a fixed value of j . Well, since j controls the y -coordinate, a fixed value of j fixes your horizontal strip.

And what you're saying is I will now add up all the masses along that strip. Then what you say is, and then as I do that, let me perform that for each j as j goes from 1 to m .

In other words, what this would say is add up all of the Δm s along each row, and then take the sum of all the possible values that the rows can take on. In other words, I'll add up these pieces, these pieces, these pieces, these pieces, then add those sums together. Notice, of course, I could have written this in the reverse order. Namely, this would say what?

First hold i constant, and for a fixed i , sum these as j goes from 1 to m . In other words, for a fixed i , sum these as j goes from 1 to m , and then sum over all the j 's. Over all the i 's. At any rate, the learning exercises will take care of making sure that you learn to manipulate this particular notation. The theory works as follows.

What we do is, assuming that the density function is continuous, at each particular little rectangle that we have here we, pick the point which has the smallest density. We picked the point which has the largest density.

We then imagine that we had replaced this little segment by the same area, but with

a material which has a constant smallest density, so that the true mass must be greater than the mass of that piece. And then we assume that this little piece was replaced by a new piece with the same shape, but whose density was constantly equal to the greatest density on this.

So this particular mass must be less than the mass of the new piece. And in that way, we squeeze the true mass between two extremes. Again, just as we did in calculus of a single variable. And again, to keep the notation going the same way, I let Δm_{ij} denote the mass of the piece that I get by replacing the density by the constant density equal to the smallest density, ρ_{ij} .

Treating that as a constant, multiplying that by the cross sectional area, Δx_i times Δy_j . I prefer to call that Δa_{ij} so we don't prejudice either the order in which we write these factors, or the particular coordinate system that one wants to use. But because I think you're more familiar with Cartesian coordinates, I will continue to write this.

But the idea is I compute the lowest, the smallest mass in other words, and underestimate in a similar way. I compute and overestimate. I then know that the true mass of my ij piece is caught between these two, therefore my total mass, which is the double sum of these over i and j must be because must be in turn caught between these two double sums.

In other words, what I have now done is caught the true mass between a lower bound and an upper bound. And again, to write this thing more symbolically so we can look at it here, what I'm saying is on the one hand, my mass must be at least as great as the value of this double sum. On the other hand, it can be no greater than the value of this double sum.

Somehow or other, the same as I did in calculus of a single variable, I would now like to put the squeeze on this as the sizes of the pieces Δx_i and Δy_j are allowed to go get arbitrarily close to 0 in size. What I hope happens is that in that limit, this difference goes to 0 so that m is caught between two equal things, and hence the true value of m must be that common value.

You see, in a manner of speaking, theoretically, nothing new is happening that didn't already happen in calculus of a single variable, but from a computational point of view, limits of single sums have now been replaced by limits of double sums, so that computationally, there is a degree of difficulty more present in our present study than there was in our study of calculus of a single variable.

The computation becomes more difficult, the theory stays the same. Now before we start getting into the idea of what limits mean, let me just summarize this for you in terms of a more concrete interpretation. Let me again take that square R with its vertices, the same one we dealt with. 00101101.

And let me now take as a special case the case in which I divide both the x region and the y region into two equal parts, so that my points of partition, you see, look like this.

Now you see, the advantage of picking my density to be $x^2 + y^2$ is that since the density is radial, what this means is that in each of my partition rectangles, in this case, each of my rectangles is a square. You see, keep in mind that first of all, I don't have to divide these two dimensions into the same number of equal parts. Secondly, they don't even have to be equal parts.

But I've, just for symmetry here, elected to do this. Just so we get an idea of what's happening computationally, and the more difficult cases will be taken care of in the exercises. The idea now is what? That for each of these particular squares, the lowest density occurs at the point which makes up the lower left hand corner, because that will be the closest point to the origin in each of these squares.

And the furthest point from the origin in each of these squares would be the point in the upper right hand corner. So you see in this particular case, let me do both of these at once. To find the smallest density in each of these regions, let me have the smallest one written in the black chalk on the lower part of this diagonal. You see the 0.00 corresponds to a density of 0.

So the lowest possible density is 0 on this block. The point $(\frac{1}{2}, 0)$ is the closest point

to the origin in this little square. So $x^2 + y^2$ in that case is simply $1/2^2 + 0^2$, which is a quarter. The closest point here is $1/2, 1/2$, so if we square the coordinates and add, we get $1/4 + 1/4$, which is $1/2$. And in a similar way, the lowest density in this block is also $1/4$.

Now again, using the same argument, the point in the first block which is further from the origin is $1/2, 1/2$. That density corresponds to $1/2$. In other words, $1/2^2 + 1/2^2$. So now putting in the biggest densities in each block, and carrying out the computations in a very trivial way. We see that the maximum densities are $1/2$ here, $5/4$ here, 2 here, and $5/4$ here.

The idea is going to be now that what we can now do is we know that the area of each of these pieces is $1/2$ by $1/2$, which is a quarter. I will now assume two different plates. One will be the plate which has the same shape as this, but in this quadrant here, has a constant density of 0 . In this quadrant, here has a constant density of $1/4$. In this quadrant here, a constant density of $1/2$.

And in this quadrant, a constant density of $1/4$. I can now compute that mass, and I know that mass must be less than the mass that I'm looking for because I have put the lowest possible density in each quadrant.

Correspondingly, if I now compute the mass of another plate which has the same shape, but in which each quadrant has the density of the greatest value of the density in each quadrant, that must give me an over approximation. And so my mass is caught between these two. And just to let you get an idea of what this notation means, noticed the $\Delta x \Delta y$ in this case is $1/4$.

For i equals $1, 2$, and j equals 1 and 2 , in other words, our elements of area are a sub $1, 1$, a sub $1, 2$, a sub $2, 1$, a sub $2, 2$, where if you want to do these in order, just let i be 1 and 2 here, j be 1 and 2 here. For example, the maximum density in the first row, second column, well, I'm reading these the wrong way here. The $\Delta x \Delta y$ First row, second column up this way is $5/4$.

The minimum density in that same quadrant is $1/4$, et cetera. At any rate, what I'm

saying is if I now take the fact that for constant density, the mass is the density times the area, a lower approximation to my mass is obtained by adding up all of my constant densities and multiplying by the area. In other words, 0 times $1/4$ plus $1/4$ times $1/4$ plus $1/2$ times $1/4$ plus $1/4$ times $1/4$.

That sum comes out to be $1/4$. Correspondingly, the upper sum is what I get by taking each of the constant densities and multiplying it by the constant area, $1/4$. That leads to a sum of $5/4$. And therefore, no matter what else I know about the mass of my region R , I know that it must be more than $1/4$, but less than $5/4$.

And the idea is that by putting on the squeeze, by making more and more subdivisions, I can hopefully find the exact value of m sub R . But finding this exact value involves a very, very difficult computation. Namely, I divide this up into m times n little elements of area. I pick a point in each region.

The density at a particular point that I've picked out in that region, which I'll call C_{ij} , D_{ij} is C_{ij} squared plus D_{ij} squared. This would give me the mass of a particular piece, and I add these up as i goes from 1 to n , as j goes from 1 to n , and take the limit as the maximum ΔX sub i , and the maximum Δy sub j approach zero. And this, in most cases is very, very difficult.

And in fact, even in this relatively simple case, what do I mean by relatively simple case? In terms of a region, what could be more straightforward than our little square r that we chose? In general, we'll have more complicated regions. Let me summarize today's lesson, essentially, by showing you in particular what it is that we're talking about in terms of double sums.

Namely, what we're going to do is we'll assume that r is a reasonable subset of two dimensional space. And by reasonable, I mean that it's not actually a kinky type of thing that leads to pathological problems. The technical words are things like r is simply connected and things like this. We'll talk about that in more detail. What I'm saying is let me imagine that r is a fairly well defined region in the xy plane.

Let me assume that I have a mapping f that carries r into e . What does that mean? f

is a function a real valued function of two independent variables, namely r is a two dimensional space. So a typical element in r is a two-tuple x comma y , and what we're saying that f maps x, y into some number. Remember e, e_1 , if you want to call it that, is the set of real numbers.

So the idea is let's, what we can say now is this. Let's take this region r . Let's partition it into a mesh of little rectangles here. Let's, in the ij rectangle, namely the rectangle whose dimensions are Δx_i by Δy_j . Let's pick a particular point, and to identify it, we'll call it c_{ij}, d_{ij} . Let's compute f at that point.

Then we'll compute f times this area, and we'll sum this thing up as i goes from 1 to n , and j goes from 1 to n, m , and we're going to take a limit as this partition gets as fine as we wish, meaning the size of the pieces get arbitrarily close to zero in area. Both in the x_n and the y direction. Notice, by the way, that because this is not a rectangular region, pieces of rectangles are left out.

And that from a computational point of view, as you make this refinement greater and greater, this is why the theory is so voluminous on this particular type of topic. You must show that you're squeezing out all the error and things like this. But conceptually, all that we're really doing here is the following. What we do is we pick a particular point in a particular one of these rectangles.

We evaluate f at that particular point, which exists because we're assuming that f is defined on this. And if you're having trouble visualizing this, f has nothing to do with this boundary. F is a function which maps each point into a number, so graphically, you can think of f as being a surface here. Namely if we visualize f of x, y as being mapped with respect to the z -axis, we have a surface coming out here. And in a sense, what we're really trying to do is to find an approximation for the volume of the region cut off by this cylinder.

And whose top is the particular surface z equals f of xy . But forgetting about that for the time being, mechanically, all we do is we from this particular double sum. We formed this double sum, and now I've put the brackets in here to emphasize that what this says is you first pick a fixed value of j . And sum this over, all i from 1 to n .

And when you're all through with that, the i no longer appears.

You have something that depends only on j . And you now do what? Sum this result up over all values of j as j goes from 1 to m . And you then compute the limit as the maximum δx_i and maximum δy_j both approach 0. Now first of all, this limit may not exist. Secondly, the limit may exist, but it depends on the order in which you're adding these terms.

If after all, we saw that we could change the sum of a simple infinite series by changing the order of the terms, certainly it shouldn't be surprising that an infinite double sum is effected, it shouldn't be surprising that is effected if we change the order of the terms. So we have one of two ways of doing this in terms of being systematized. Do we first sum in terms of the x 's, and then add up the y increments?

Or the other way around? The key result is this, though. It's that this particular limit will exist, and will be independent of what order you add these things in, provided only that the function f is at least piecewise continuous. In other words, that f is a continuous function. And if it's not continuous, it has breaks of only a finite number of places.

And the idea is that not only will this limit exist, but in the manner analogous to that definite integral. When this limit exists, we denote it by what? We replace this by x and y . This gets replaced by dx and dy , and the double sum gets replaced by a double integral, and what we now do is indicate what the region r is over here.

In other words, in terms of a summary, whenever I write down this expression, it's an abbreviation for a limit of a double sum. And what limit of a double sum is it? It's the particular limit of this double sum if it exists, and that limit will exist if f is piecewise continuous.

Now, the key point that I want to bring out here, and in fact, this will lead into the lecture of next time, also, is that if I had never heard of partial derivatives, it makes sense to talk about this kind of a sum. In other words, note that forming this limit

does not require any knowledge of partial derivatives.

In the same way that finding infinite sums in calculus one could be done independently of any knowledge of a derivative, this particular concept, this particular limits makes sense even if we've never heard of partial derivatives. And what does this double sum represent? Just by way of a quick summary again, there are two very specific physical interpretations we can give this.

One is that this double sum represents the density of the plate whose shape is the region r . The mass of the plate whose shape is the region r , and whose density at the point x, y and r is given by f of xy . In other words, this part here controls the density, and this part here controls the shape of the region.

A second interpretation is that this limit, when it exists, represents the volume of the right cylinder. That means a cylinder is obtained by tracing all over the region r with a line perpendicular to the xy plane. That particular cylinder whose top is the surface z equals f of xy . Here are, then, the two interpretations that we give this.

One is it can be a mass. The other is it can be a volume.

The key point is that to evaluate these double sums, these limits of double sums, so if you want to call them infinite double sums, there is no necessity for knowing partial derivatives. However, one might expect that in the same way that there was a connection between single infinite sums and ordinary antiderivatives, there should be a connection between double or multiple sums and anti derivatives of partial--

Well, the antiderivative involving partial derivatives. It turns out that this is indeed the case. We will talk about this in more detail next time, but between now and next time, what I would like you to do is to drill particularly hard on these exercises, and make sure that you become familiar and feel at home with the notation that's used in denoting double and other multiple sums. At any rate then, until next time, good bye.

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.