Basic Calculus - 1

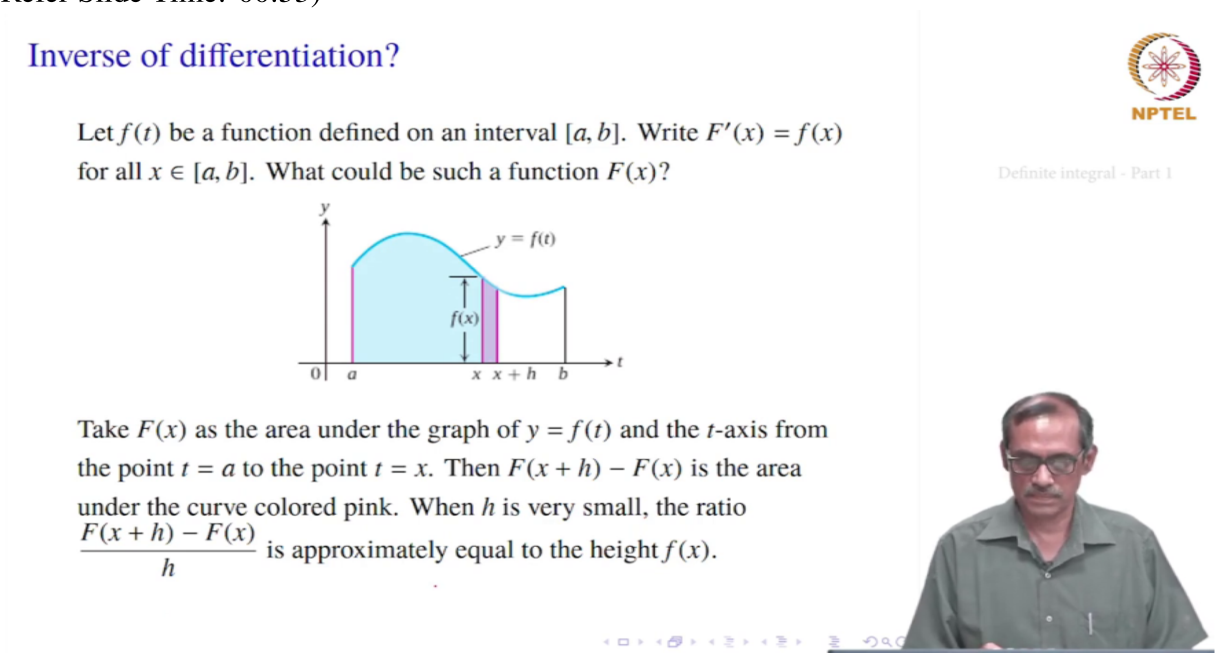
Professor. Arindama Singh
Department of Mathematics
Indian Institute of Technology Madras

Lecture 25 - Part 1

Definite Integral - Part 1

This is lecture 25 of basic calculus 1. Till now we are concentrating on very important applications of the notion of limiting concept. Limiting concept is of course, the most fundamental thing to Calculus. Using this we have discussed continuity and then differentiation of functions. Today we will be starting the third important aspect which also bases on the limiting notion; it is called integration.

(Refer Slide Time: 00:53)

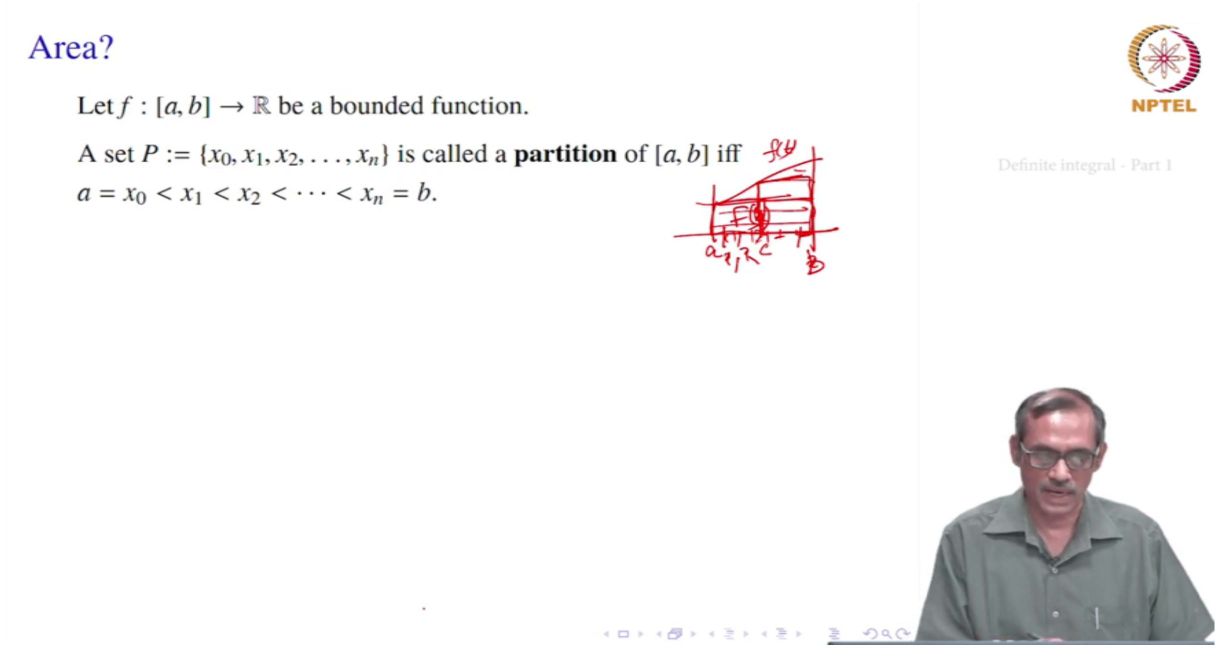


As it is told, integration is the inverse process or reverse process of differentiation. This is the sense. Suppose $f(t)$ is a function defined on the closed interval $[a, b]$. If we can write $f(x)$ equal to derivative of some other function, say, $F(x)$, then what could be this new function $F(x)$? This is how it is called the reverse process of differentiation. In differentiation $F(x)$ is given, you differentiate and obtain something as small $f(x)$. Here we are given $f(x)$ and ask "what should be $F(x)$ such that the derivative of $F(x)$ will be giving us $f(x)$?"; that is the idea.

So, x is any point in the closed interval $[a, b]$. The functions $f(x)$, $F(x)$ and $F'(x)$, the derivative of $F(x)$ should be defined there, and $F'(x)$ should be equal to $f(x)$. Look at the picture. Suppose $f(x)$ is given. Let us say it is the t-axis, and we have $y = f(t)$. We choose any point x here. We want to find $F(x)$. where this height is $f(x)$. Now, what do we guess? Suppose we consider $F(x)$ as the area under the graph of $y = f(t)$, the t-axis, and the lines $t = a$ and $t = x$. It is this area which is painted blue here. Suppose that is equal to $f(x)$. Then $f(x + h)$ will be equal to that plus this pink area. Now, that is $F(x + h)$. If you take $F(x + h) - F(x)$, that will give this pink area. Think of this area divided by this h. It is some average value of the heights. And when this h becomes smaller and smaller, we may say that we reach the height $f(x)$. That is the intuitive notion behind getting this function $F(x)$.

So, we guess that $F(x)$ should be the area under the graph of $y = f(t)$ and the lines $t = a$ and $t = x$, which is painted blue here. That area should be $F(x)$ because that matches with our intuition. If you take $f(x+h) - f(x)$ divided by h, then that is approximately equal to the height, which is the value $f(x)$.

(Refer Slide Time: 4:32)



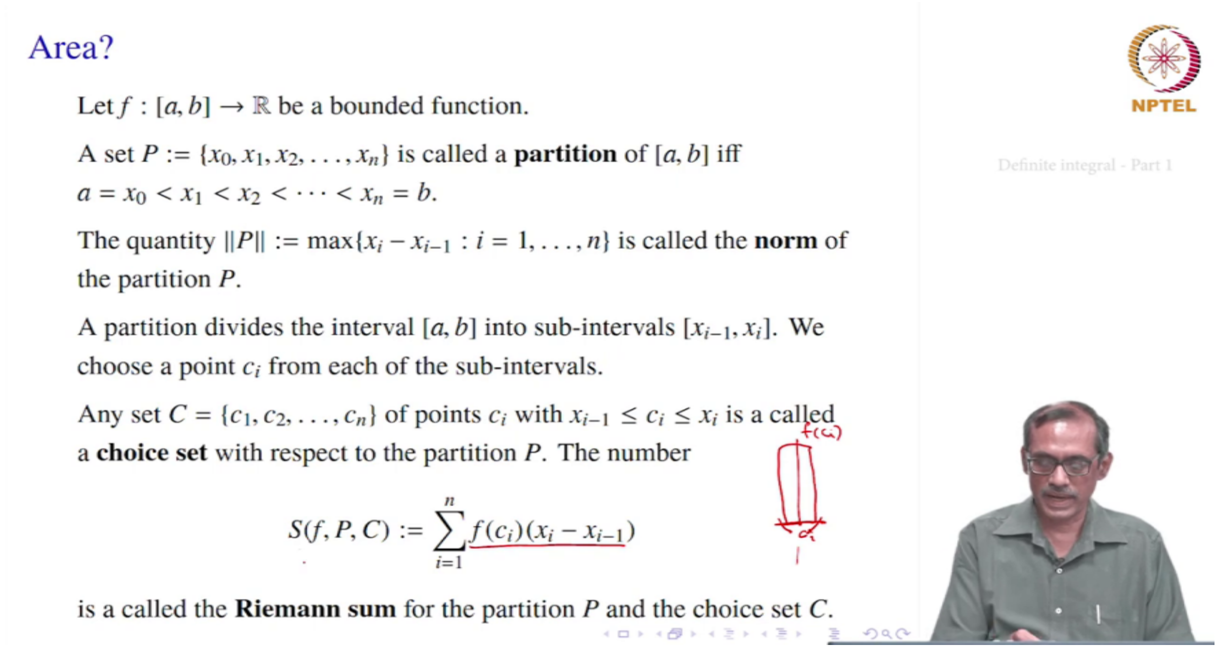
Now, a new question arises. How to find the area of this region? Suppose we say that there is some area. Then first of all f must be a bounded function. If it is unbounded, then the area becomes ∞, right? That will not help us; it will not be a real number, and we cannot write $f(x)$ since that is ∞. So, we assume that $f(x)$ is a bounded function. But we are trying to find the function $F(x)$. In fact, we will put some further restrictions, because in applications we get usually that kind of functions; so, that will be sufficient for our purpose.

Suppose $f : [a, b] \to \mathbb{R}$ is a bounded function. In the picture, it is a and this is h and that is our $f(x)$. The problem is, how to approximate this area under the curve $y = f(t)$ bounded by these two lines? How do we approximate this area? Instead of x, we can take b; that does not matter if that fixes the idea. How do we approximate this area if we cannot find it exactly?

What do we guess? If we take this rectangle from a to b, then that is also an approximation of this area. But a better approximation will be obtained if we divide that into two sub intervals and now take the sum of these two areas. It is some c, so that we get one rectangle with a to c and another from c to b. We may think of approximating the area under the curve by the sum of these two. Intuitively, if we go on subdividing this interval $[a, b]$ with smaller and smaller sub-intervals, then on each of the sub-intervals we compute that rectangle, and add them. That will be a better approximation when these sub-intervals are smaller and smaller. That is the idea.

To capture this notion of sub-intervals, we have to choose some points here. The set of those points is called a partition. So, we are defining it formally now. We should try looking at the partition itself. A partition of $[a, b]$ is set P which is $\{x_0, x_1, x_2, \ldots, x_n\}$, where $a = x_0$, $b = x_n$ and other points are chosen between a and b with this ordering. We have $x_n - x_0 = b - a$. That means, you have a which is x_0, then x_1, then x_2 and so on so that x_n is b. After this, we have to do something with this n. We want the sub-intervals to be smaller and smaller. Then this would introduce a limiting process so that we may get this area.

(Refer Slide Time: 7:44)



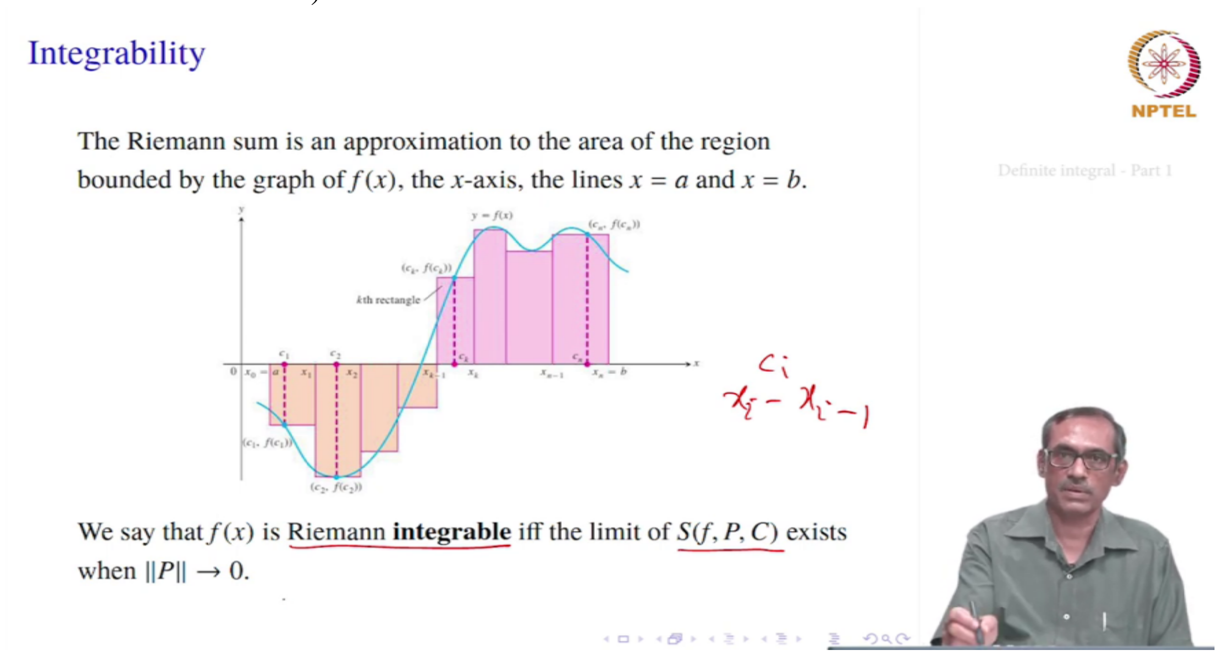
Since we want the sub-intervals to be smaller and smaller, and in fact their lengths should be close to 0, we define the norm of the partition. The norm of P is the maximum of the lengths of those sub-intervals. If that maximum goes to 0, then the lengths of all these sub-intervals will also go to 0. That is why we are defining this norm. So, let us call $\|P\|$ as the norm of the partition P, which is really the maximum of $x_i - x_{i-1}$. We see that this partition divides the interval $[a, b]$. And where from a rectangle will come within this sub-interval $[x_{i-1}, x_i]$? You may take the area of this rectangle as $f(x_i)\,(x_i - x_{i-1})$. In general, we will choose another point c_i in between x_{i-1} and x_i and take the area $f(c_i)\,(x_i - x_{i-1})$.

This rectangle obtained by choosing a point c_i in between x_{i-1} and x_i is the area of the rectangle limited to this sub-interval $[x_{i-1}, x_i]$. We will say that this area of the rectangle somehow approximates the area under the curve $y = f(t)$ bounded by the lines $t = x_{i-1}$ and $t = x_i$.

We will give a name to the set of all these chosen c_is. We will call it a choice set. This is non-standard; but it will be very helpful to fix the notion. Let us call this set $C = \{c_1, c_2, \ldots, c_n\}$ as a choice set. A choice set consists of points from each of the sub-intervals; that is, c_i should be in between x_{i-1} and x_i. Then, we form the sum of all those rectangles. The area of the rectangle in the interval $[x_{i-1}, x_i]$ is $f(c_i)$, which is this height at c_i times the length of the sub-interval, which is $x_i - x_{i-1}$. Then, you take the sum of all these things. This sum is thought to be an approximation to the area under the curve. Let us give a notation. We call this sum the Riemann sum of f for the partition P and the choice set C; and we write it as $S(f, P, C)$.

This $S(f, P, C)$ will depend on the function f, the particular partition P, which is defined by the particular points x_i, the break points that make the sub-intervals and it also depends on the choice points. It is really the sum of the areas of the smaller rectangles, and we call it the Riemann sum. Our idea is, this is an approximation to the area, and this will be equal to the area when each of these lengths $x_i - x_{i-1}$ goes to 0. In that case, the norm of P will go to 0, since it is the maximum of those lengths. And when $\|P\|$ goes to 0, each of these lengths will also go to 0. And that is exactly our definition of the area.

(Refer Slide Time: 11:38)



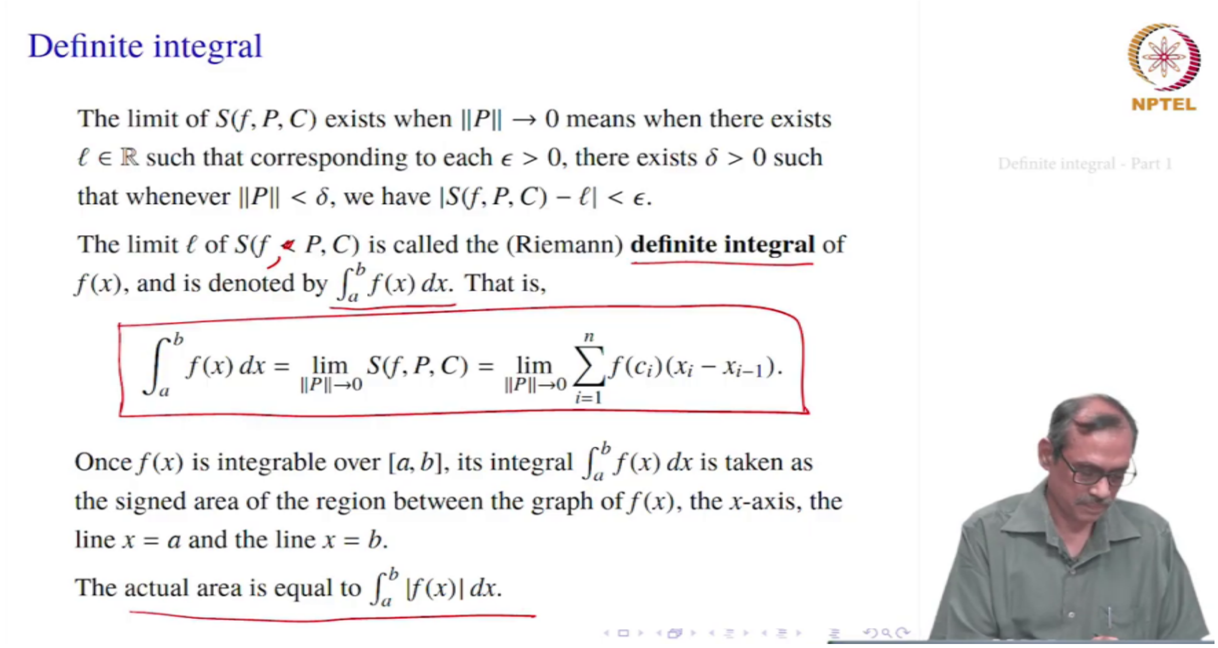
Let us look at the picture to fix the idea. You have the interval $[a, b]$. You get the sub-intervals by taking these points of the partition. Inside each of these sub-intervals, you have chosen points c_i and then found out the area of the rectangle with the height as c_i and base as $x_i - x_{i-1}$; then you take the sum of all those areas of rectangles; and you think of that sum as an approximation to the area bounded by the curve $y = f(x)$, the x-axis and the lines $x = a$ and $x = b$.

See, it might happen like this, because it need not always be lying above the x-axis, the curve can be below the x-axis. So, when you take this region, it might give rise to two different regions and we are taking their sum. Now, what is the area of the region bounded by the curve, the x-axis and those two lines? That will be the limit of this Riemann sum $S(f, P, C)$ as $\|P\|$ goes to 0, provided it exists. There can be some curves where area does not exist. We are assuming that possibility here.

We will give it a name. Whenever this limit exists, we say that the function $f(x)$ is Riemann integrable or just integrable. Here, the limit of the Riemann sum should exist when $\|P\|$ goes to 0. When $\|P\| \to 0$, it will mean that each length of these sub-intervals, that is, $x_i - x_{i-1}$ will become smaller and smaller and it will approach 0. So, it assumes that whatever choice point c_i you choose in between x_{i-1} and x_i, the limits would exist.

For each choice set you will get an approximation. But whatever choice points you choose, when the norm of the partition goes to 0, this limit must exist. That is the condition of Riemann integrability. If that limit exists, we say that the function is Riemann integrable. And we will identify that limit with the area. That is exactly our definition.

(Refer Slide Time 14:16)



We say that if the limit of $S(f, P, C)$ exists as $\|P\| \to 0$, then that value of the limit will be the area. What is the meaning of this limit exists when norm P goes to 0? It simply means that the limit must be a real number ℓ. Then, this limit exists means the absolute value of the difference of this Riemann sum from ℓ can be made as small as possible by choosing our partition P with the condition that $\|P\|$ goes to 0.

Abstractly, we can define it this way: there exists an ℓ, which will be the limit of that Riemann sum, if corresponding to each $\epsilon > 0$, there exists a $\delta > 0$ such that whenever $||P|| < \delta$, we should see that the difference between these two quantities is less than ϵ. It is just like the definition of limit of functions defined earlier.

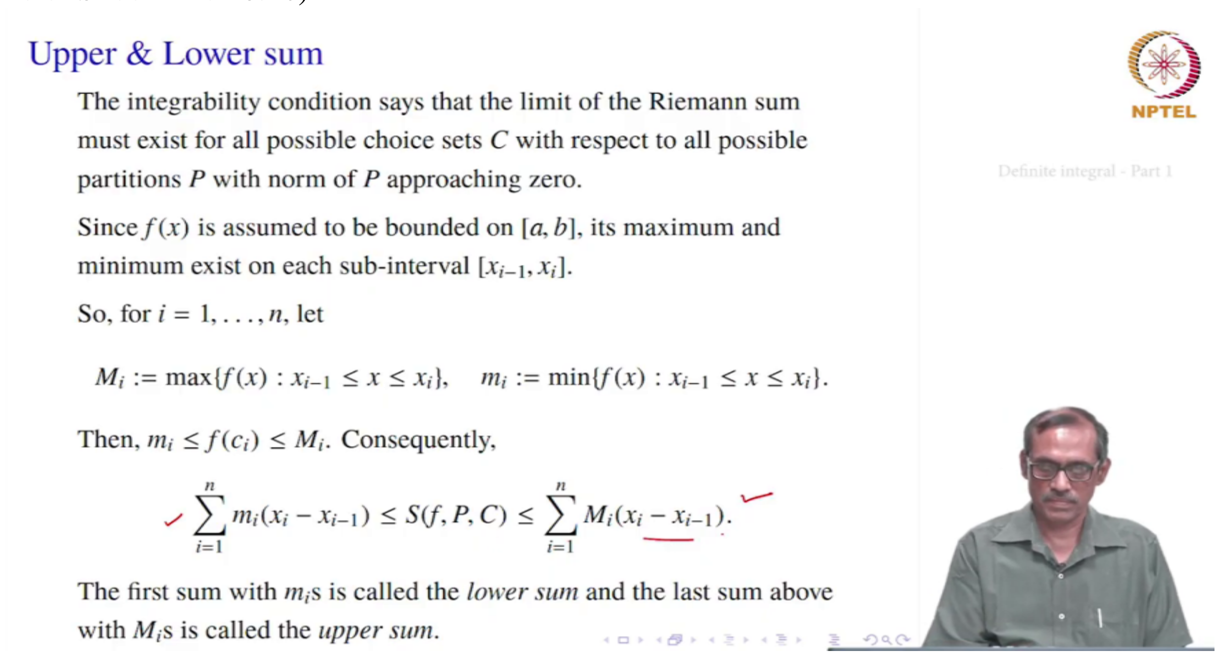
If this condition is satisfied, we would say that the function f is integrable. Of course, its integral will be that ℓ. We will give a name to this ℓ. We will call that as the definite integral or the Riemann integral of $f(x)$. And we will denote that limit as $\int_a^b f(x)\,dx$. The integral is from a to b, and we write these $f(x)\,dx$ as a notation. The definite integral is equal to the limit of the Riemann sum $S(f, P, C)$ when $\|P\| \to 0$. It is the same thing as telling that $\int_a^b f(x)\,dx = \lim_{\|P\|\to 0} \sum_{i=1}^{n} f(c_i)(x_i - x_{i-1})$.

This will be our definition of the integral, which is actually the area under the curve $= f(x)$, the lines $x = a$, $x = b$, and the x-axis. In fact, this is the signed area. Why signed? You can see from the previous picture where $f(x)$ is negative right up to this point. All these areas will give you (not exactly areas in our geometric sense) negative sign. Suppose this length is 2 and this height is 3.

Then you get -6 because this height is $f(c_i) = -3$. So, you get -6; thus it is really the signed area; it is not exactly the area. That means, we have defined the signed area of the region between the graph of $y = f(x)$, the x-axis and the lines $x = a$ and $x = b$.

If you want to get the correct or exact area, you have to take $\int_a^b |f(x)\, dx$; that will give us the geometric area. That is what the actual area is; it is equal to the integral of $|f(x)|$ from a to b. The logic is that the earlier negatives will become positive now so that we get the actual area.

(Refer Slide Time 18:10)

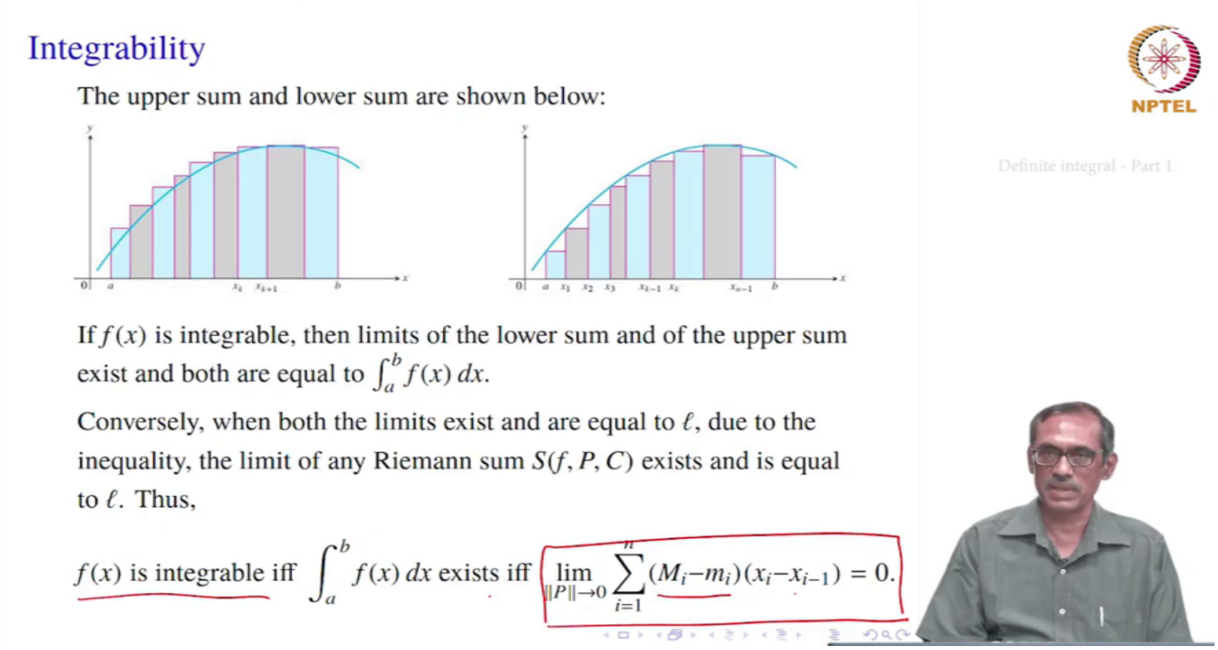


Once this definite integral is defined, we should see how it is applied. And then slowly we will come to some other conditions which will imply that the integral exists. It is easy to see what should be the conditions. You see that the limit of the Riemann sum exists for all possible choice of c, that is what our result is. With whatever partition you choose, once the norm of the partition goes to 0, whatever choice set C you choose, it should give us a limit.

Now, our $f(x)$ was assumed to be bounded. Therefore, in each of these sub-intervals, $f(x)$ is again bounded. So, it has a maximum and it has a minimum. Then, in this sub-interval $[x_{i-1}, x_i]$, the term $f(c_i)(x_i - x_{i-1})$ will lie between the maximum of $f(c_i)$ times $(x_i - x_{i-1})$ and the minimum of $f(c_i)$ times $(x_i - x_{i-1})$. Let us write $M_i = \max f(x)$ and $m_i = \min f(x)$ where $x \in [x_i - x_{i-1}]$. Then, $m_i \le f(c_i) \le M_i$. Therefore, the Riemann sum $S(f, P, C)$ lies between the other two sums where we replace $f(c_i)$ with m_i and M_i.

Notice that if the minimum is achieved somewhere, say, at some t_i, then you can choose this c_i as t_i. In that case, the Riemann sum $S(f, P, C)$ will be the smallest among all Riemann sums. Similarly, if the maximum is achieved at some points, we can choose the points c_i to be those points. And then the Riemann sum will be the maximum of all Riemann sums. We will give some names to these two particular sums. We will call this left side as the lower sum and the right side as the upper sum. So, the Rieman sum lies betwen these two, the loer sum and the upper sum. In particular you can choose c_is so that the Rieman sum is equal to the lower sum, and you can choose

possibly different c_is so that the Riemann sum is equal to the upper sum.
(Refer Slide Time 21:08)



So, it may look something like this. Suppose inside the interval x_{i-1} to x_i, the curve is going like this. Your x_i is a point where this $f(c_i)$ is maximum, as in this picture. When you come to the other side, for example here, it is maximum at this place not at this place. So, if c_i is chosen between these two, then its maximum will be at the left point. Well, whether left or right, we do not know, but there is some point where it is maximum. Here you see on the other picture that the c_i is taken to be minimum. Then, you get these rectangular areas while you computing the minimum.

So, the lower sum is $\sum_{i=1}^{n} m_i(x_i - x_{i-1}$ and the upper sum is $\sum_{i=1}^{n} M_i(x_i - x_{i-1}$. These m_i and M_i may correspond to some particular choices of c_i. When $f(x)$ is integrable, this lower sum and the upper sum should exist. Because for every choice of c_i the limit should exist; and these are particular choices. When you take the limit, both must be equal to the integral $\int_a^b f(x)\,dx$. So, when both the limits exist and they are equal due to our inequality that the Riemann sum lies between the lower and upper sums, the Riemann sum in the limit must exist and that should be equal to this integral $\int_a^b f(x)\,dx$, which is ℓ. Fine. That would say that $f(x)$ is integrable if and only this integral exists, if and only if the limit of $S(f, P, C)$ is equal to ℓ, if and only both the lower sum and upper sums give the same limit.

The last condition of integrability means: the lower sum and the upper sum have the same limit ℓ. It can be written in a better way without bringing in this ℓ. It is simply the condition that the limit of their difference is 0. That is, the limit of $\sum_{i=1}^{n}(M_i - m_i)(x_i - x_{i-1})$ as $\|P\| \to 0$ must be 0. Sometimes if we are not able to compute the integral correctly, we can at least show by using this method that the function is integrable.