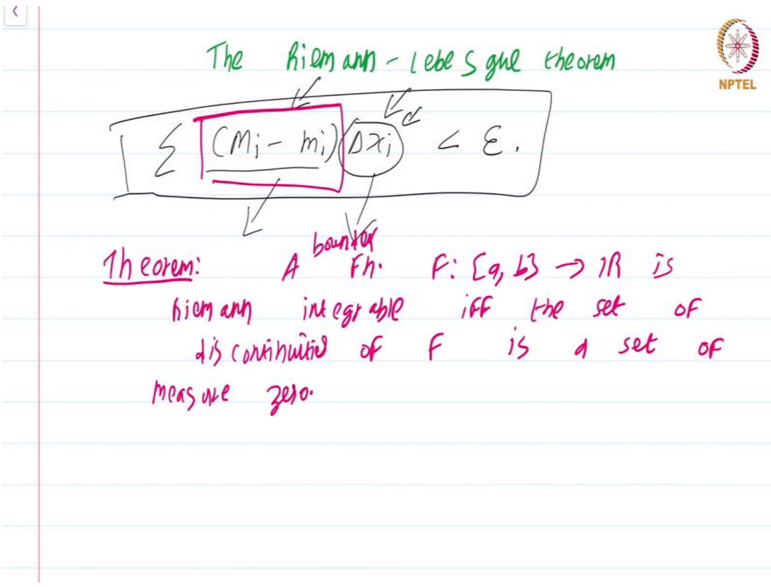
Real Analysis - I
Dr. Jaikrishnan J
Department of Mathematics
Indian Institute of Technology, Palakkad

Lecture – 27.2
The Riemann-Lebesgue Theorem

(Refer Slide Time: 00:14)



We now come to the Riemann-Lebesgue Theorem. We are going to completely characterize precisely which functions are continuous. Then we are going to get a huge list of corollaries, pretty much all the major theorems of Riemann integrals will follow from this theorem. Now, in traditional treatments at the undergraduate level this theorem is usually skipped.

And instead these properties are proved directly by using the condition that a function is integrable if and only if $U(f, p) - L(f, p)$ can be made less than ε. Now, the reason why I have preferred to prove the Riemann-Lebesgue theorem, whose proof by the way is quite challenging, is the following. There is a deep idea behind the proof and as we all know ideas are bulletproof.

This idea in various forms is there in the proof of the various properties. By centralizing all the deep ideas into one theorem and understanding that once and for all, we have economy of thought. You just have to understand this theorem and everything about Riemann integration will follow from this.

So, what is this deep idea that I am talking about? It is essentially this sum, $\sum(M_i - m_i)\Delta x_i$; we want to make this small. For each ε, if you can find a partition such that this is small then the function f is integrable and vice versa. Now, there are two obstacles to this sum being small. One is that this could be really large. This $\sum(M_i - m_i)$ could be really large, second is that this could be really large.

But, notice that we have full control over this Δx_i , right. We choose the partition. The statement of the characterization of Riemann integrability is that, given any $\varepsilon > 0$, we can find a partition P such that $\sum(M_i - m_i)\Delta x_i < \varepsilon$. We have full control over this Δx_i. Therefore, what we are going to do is we are going to delicately balance this quantity which could be large at some places and this quantity for which we have full control.

So, the essential idea is the following. Let me first state the theorem; so, that I am not thinking out aloud in air. Let me have something concrete in front of you and me. Theorem: A function $f: [a, b] \longrightarrow R$; let me add a bounded function of course, a bounded function is Riemann integrable if and only if the set of discontinuities of f is a set of measure zero.

Now, away from the set of discontinuities, we have control over this $(M_i - m_i)$. We can make it as small as we desire by shrinking the intervals, right. The key is at those points of discontinuity, where the function oscillation is going to be greater than 0 we control $\sum(M_i - m_i)\Delta x_i$ by making Δx_i really small.

And, we can do this because the set of discontinuities is subset of measure zero. Therefore, you can cover it by as small open intervals as you desire. So, that is the key fact.

(Refer Slide Time: 04:08)

Theorem: A fn. $f: [a,b] \to \mathbb{R}$ is Riemann integrable iff the set of discontinuities of f is a set of measure zero.

Proof: Let us assume f is integrable. The set of discontinuities

$$D = \bigcup_{k=1}^{\infty} D_k$$

$$D_k = \{x \in [a,b] : osc_x(f) \geq \frac{1}{k}\}.$$

So, let us begin the proof and please keep in mind these vague remarks that I am making, you shall see you will see these vague remarks coming in front of you alive. So, let us assume first f is integrable ok. The set of discontinuities D; let us call it D is actually going to be equal to, you can break it up as, we have done this before as $\cup_{k=1}^{\infty} D_k$, where D_k is just the collection of $\{x \in [0,1]: osc_x(f) \geq \frac{1}{k}\}$,ok. Look at the set D_k which is the set of points $x \in [a, b]$, look at the set of points where the oscillation is greater than or equal to $\frac{1}{k}$, call that set D_k, D naturally decomposes as the union of the various D_k's ok.

(Refer Slide Time: 05:25)

Proof: Let us assume f is integrable. The set of discontinuities

$$D = \bigcup_{k=1}^{\infty} D_k$$

$$D_k = \{x \in [a,b] : osc_x(f) \geq \frac{1}{k}\}.$$

Sufficies to show each D_k is a set of measure zero. Fix $k \in \mathbb{N}$, and $\varepsilon > 0$.

Let P be a partition such that

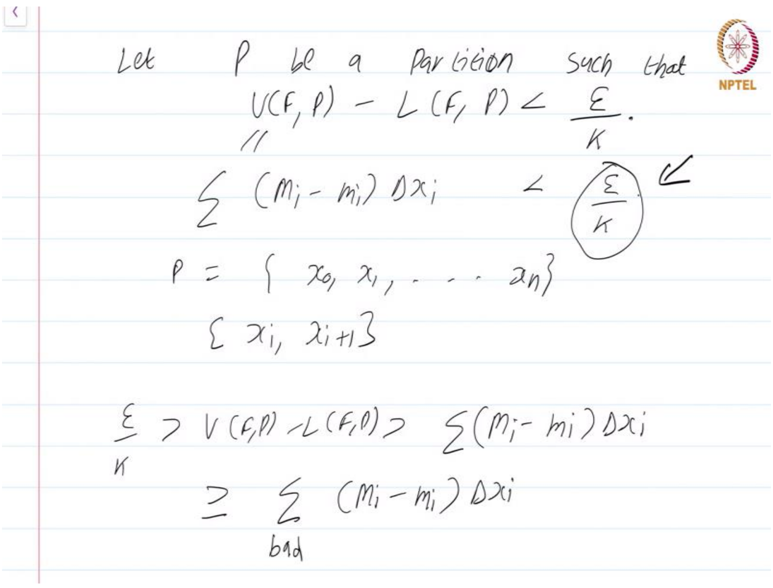
$$U(f,P) - L(f,P) < \frac{\varepsilon}{k}.$$

Now, suffices to show each D_k is a set of measure zero. Recall from the previous module that a countable union of measure zero sets is a set of measure zero. This follows by $\frac{\varepsilon}{2^n}$ trick that we have already seen ok. We are going to show that each D_k is a set of measure zero.

So, what we do is the following, let P be a partition such that $U(f,p) - L(f,p) < \frac{\varepsilon}{k}$. You will understand why I am putting this k in the denominator. This k is the same k here. We are going to show that each D_k is a set of measure zero. So, I should say fix $k \in N$ to be concrete; fix $k \in N$.

I am going to show that D_k is a set of measure zero, Fix $k \in N$ and $\varepsilon > 0$. We can find a partition such that $U(f,p) - L(f,p) < \frac{\varepsilon}{k}$. This is because we are assuming that the function f is Riemann integrable, ok.

(Refer Slide Time: 07:10)

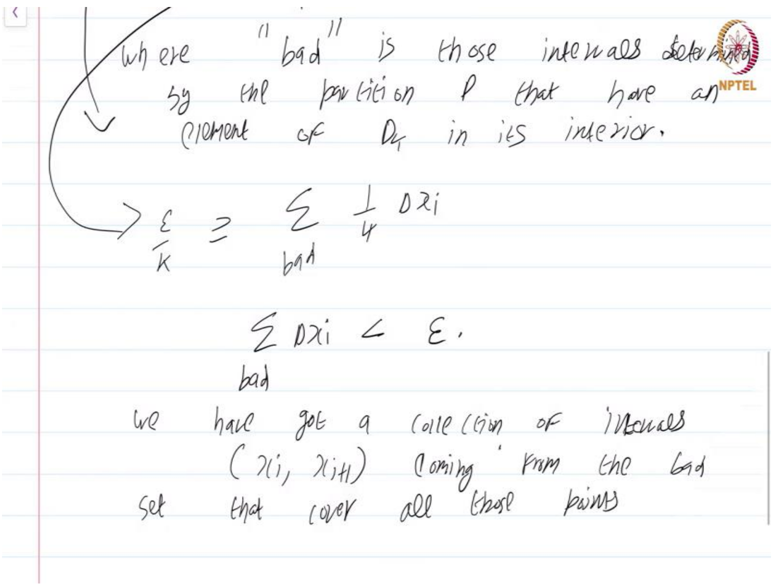


How does this help? What is $U(f,p) - L(f,p)$? It is just $\sum (M_i - m_i)\Delta x_i$. This is less than $\frac{\varepsilon}{k}$ right. How does this help? Well, observe the following because this is less than $\frac{\varepsilon}{k}$, there cannot be too many interval coming from the partition.

So, let us make that precise; let us say $P = \{x_0, x_1, \ldots, x_n\}$. There cannot be too many intervals $[x_i, x_{i+1}]$ such that the terms$(M_i - m_i)\Delta x_i$ is too large; because we have controlled it by saying that the net sum is less than $\frac{\varepsilon}{k}$.

To make this precise, we already know that $\frac{\varepsilon}{k} > U(f,p) - L(f,p) \geq \sum(M_i - m_i)\Delta x_i$, where I am going to sum up over bad. What do I mean by sum up over bad?

(Refer Slide Time: 08:34)

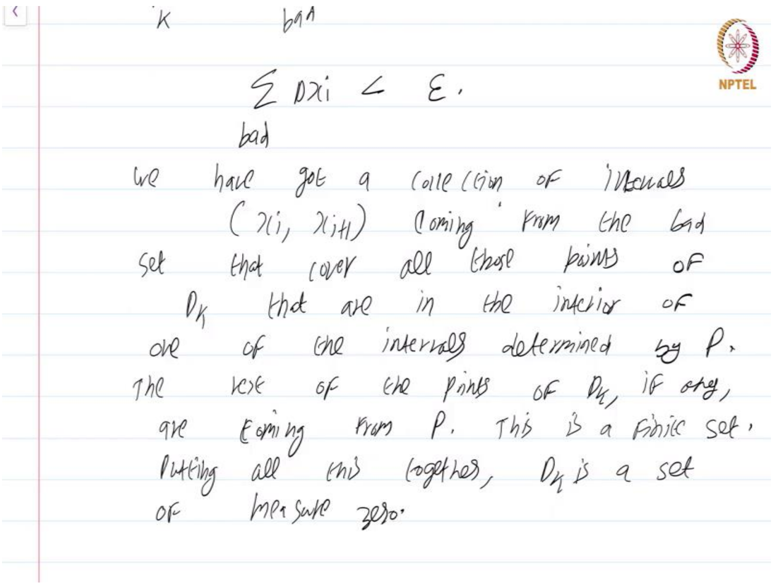


Where "bad"; let me put this in quotes, is those intervals those determined by the partition of course, those intervals determined by the partition P that have an element an element of D_k in its interior ok. So, essentially if you sum up over all these intervals, you will get the entire sum, $U(f,p) - L(f,p)$.

What I am doing is I am only summing up over those intervals such that there is a point in the interior, where there is an element of D_k; that means, a point where the oscillation is greater than or equal to $\frac{1}{k}$. Now, why am I doing this? Well, because now this term will be greater than or equal $\sum_{bad} \frac{1}{k} \Delta x_i$ right.

Now, how does this help? Well, we already know that the original quantity was $\frac{\varepsilon}{k}$, right. So, what we get is $\sum_{bad} \Delta x_i < \varepsilon$, right. Seems like a miracle has happened, but nothing much has happened; all that has happened is k have gotten cancelled ok. So, what we have got is a collection of intervals $[x_i, x_{i+1}]$; where i comes from the bad set,

(Refer Slide Time: 10:55)

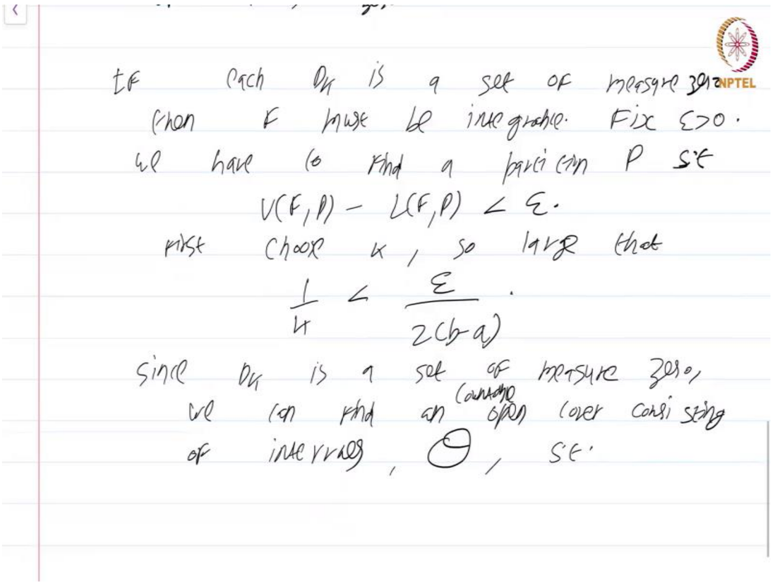


I will just instead of belaboring the point, I will just say coming from the bad set that cover all those points of D_k that are in the interior of one of the intervals determined by P ok.

So, what about the rest of the points of D_k? Well, the rest of the points of D_k, if any there might be none; if any are coming from P, right. Either a point is going to be the interiors of one of the intervals determined by this partition or it is going to be an end point. The end points are precisely the terms and the elements of the partition P right. This is a finite set.

So, putting all this together, D_k is a set of measure zero ok. So, the crux of this part of the proof was just that we already have good control over the behavior of $U(f,p) - L(f,p)$. So, the collection of intervals that happen to have a point of D_k in its interior, the lengths of those intervals cannot be too large; because we have already controlled for it by making $U(f,p) - L(f,p) < \frac{\varepsilon}{k}$. Excellent.

(Refer Slide Time: 12:39)



Now, on to part 2 which is to show that if each D_k is a set of measure zero, then f must be integrable ok; we have to show this. So, again we are going to use the criterion for Riemann integrability that there is a partition P, such that $U(f,p) - L(f,p) < \varepsilon$. If I can find such a partition then the function is Riemann integrable.

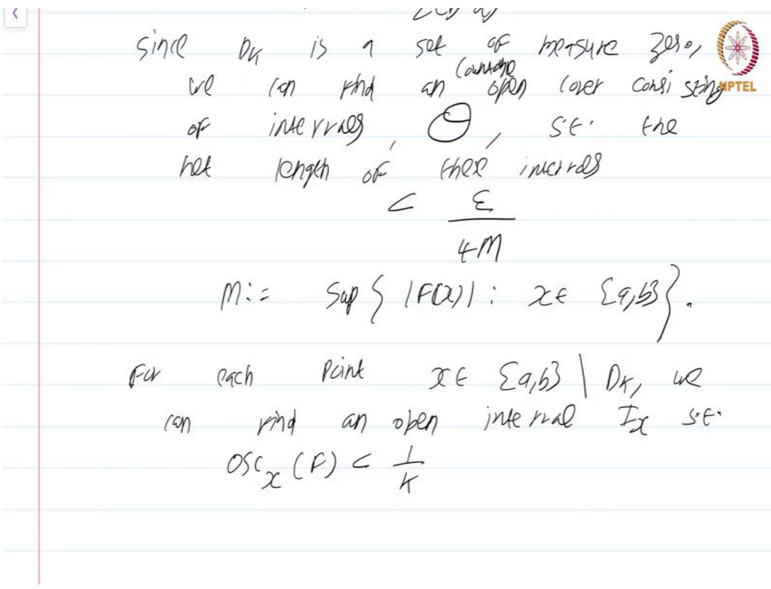
So, fix $\varepsilon > 0$. We have to find a partition P such that $U(f,p) - L(f,p) < \varepsilon$. So, now the crux is we have some control over the discontinuities. We have some control over the discontinuities, away from the discontinuities the function is nice, it behaves nicely.

So, we are going to do a proof by cases, but proof by cases in a subtle way. We are going to deal with various points depending on whether it is a point of continuity or a point of discontinuity; depending on what it is we are going to deal with it in a different manner. So, it is sort of a subtle proof by cases. So, what we do is the following; first each D_k is a set of measure zero.

So, first choose k that $\frac{1}{k} < \frac{\varepsilon}{2(b-a)}$. Now, I do not like such magic constants, but here it is sort of unavoidable. As it is the proof is a bit technical, you will understand why $2(b-a)$ in a moment ok. So, I am going to first choose k so large that $\frac{1}{k} < \frac{\varepsilon}{2(b-a)}$ ok.

Now, since D_k is a set of measure zero is a set of measure zero, I am going to pull another magic constant out of a hat; we can find we can find an open cover, countable open cover, consisting of intervals, l,et us call this fancy $\mathcal{O}$, such that the net length; the net length of these intervals of these intervals is less than $\frac{\varepsilon}{4M}$, ok.

(Refer Slide Time: 15:14)

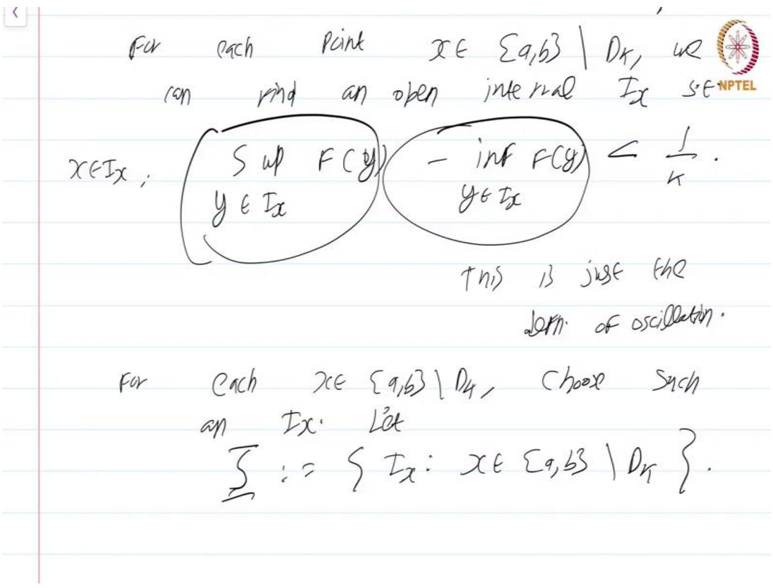


What is this capital M? Where, capital $M \coloneqq \sup\{\,|f(x)|: x \in [a,b]\}$, such a supremum exists because the function f is a bounded function ok.

So, I am going to take the maximum modulus value of this function f and I am going to choose an open cover of D_k, such that the net length of those intervals is less than $\frac{\varepsilon}{4M}$, ok. Again another magic constant, but bear with me for a few minutes, maybe 10, 15 minutes and you will understand where these constants are coming from ok.

Now, here comes the crux of the proof. What we are going to do now is we are going to control the behavior of the function away from the set D_k by using continuity. The way we are doing this is the following. For each point $x \in [a,b]/D_k$, we can find an open interval I_x such that $osc_x(f) < \frac{1}{k}$ ok; such that rather we already know this because that is how x was chosen.

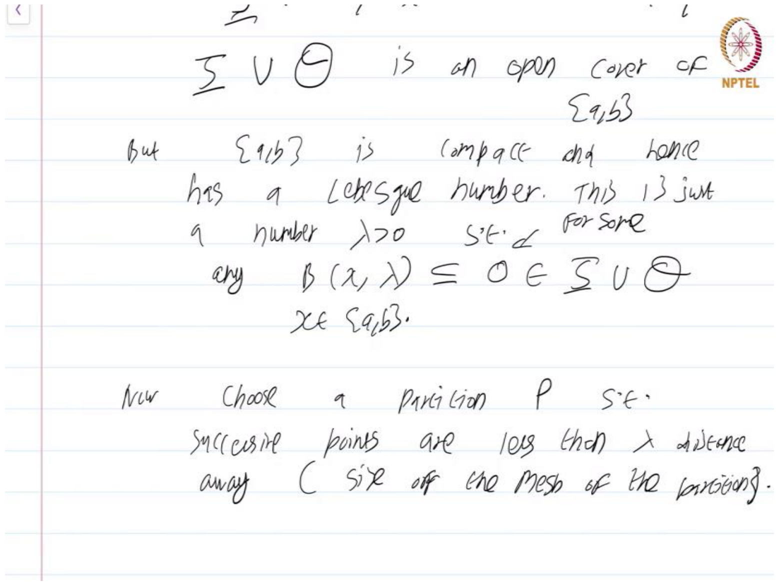
(Refer Slide Time: 17:05)



We can find an interval I_x such that $\sup_{y \in I_x} f(y) - \inf_{y \in I_x} f(y) < \frac{1}{k}$, ok. Because, you are taking a point x which is outside of D_k, the oscillation at that point has to be less than $\frac{1}{k}$; which means you can find an interval small interval I_x that contains x of course, such that x is an element of I_x.

That that interval I_x contains x and moreover when you take the supremum minus the infimum is less than $\frac{1}{k}$. This just comes, this is just the definition of oscillation, just the definition of oscillation ok. Now, we can do this for each point $x \in [a,b]/D_k$ ok. So, for each, just a second, for each $x \in [a,b]/D_k$ choose such an I_x. Let me call it some fancy $\mathcal{J}$, be just by definition $\mathcal{J} \coloneqq \{I_x : x \in [a,b]/D_k\}$.

(Refer Slide Time: 19:03)



Now, here is the thing. This fancy $\mathcal{I} \cup \mathcal{O}$ is an open cover is an open cover of $[a, b]$ right; that is because this fancy $\mathcal{I}$ covers all those points in closed interval $[a, b]$ that are not in D_k and D_k is already covered by the intervals which are in $\mathcal{O}$ right. So, this $\mathcal{I} \cup \mathcal{O}$ is an open cover of $[a, b]$. But, close interval $[a, b]$ is compact and hence and hence has a Lebesgue number.

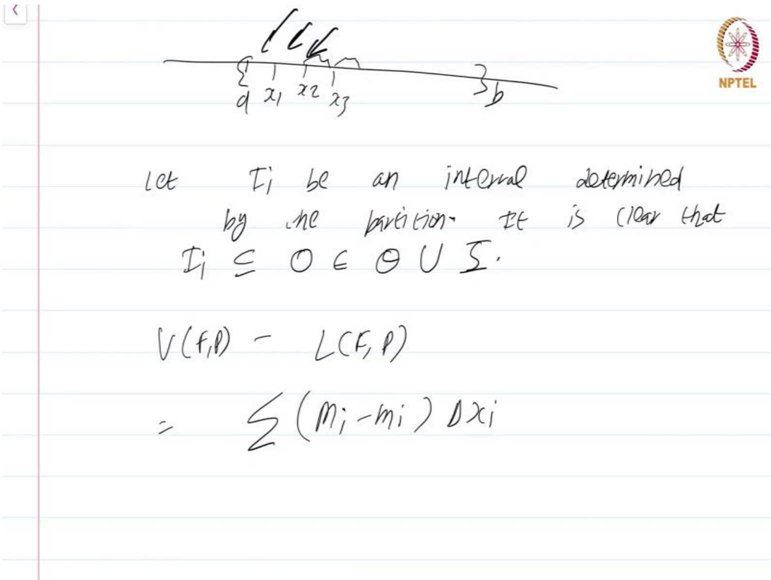
So, recall from our extensive study of topology that there will be a Lebesgue number. What does this mean? This just means this is just a number λ >0 such that any $B(x, \lambda) \subset \mathcal{O} \in \mathcal{I} \cup \mathcal{O}$ ok. So, of course $x \in [a, b]$.

A Lebesgue number is just the number λ such that if you take the ball of radius λ centered at x; where x comes from $[a, b]$. This is going to be contained in some one or more of the elements of the open cover. So, there will be some $O \in \mathcal{I} \cup \mathcal{O}$ which contains this ok; that is why that is the definition of Lebesgue number ok.

Now, here comes the crux. Now, choose a partition P such that successive points are less than lambda distance away ok. So, this is called the size or the mesh of the partition that is the least that is the least successive distance.

So, sorry not the least the largest successive distances determined by the partition, that is called the size or the mesh; it is good to know this technical term, but it is not very, it is not going to play a huge role other than just being a terminology ok. So, what we are doing is the following.

(Refer Slide Time: 21:48)

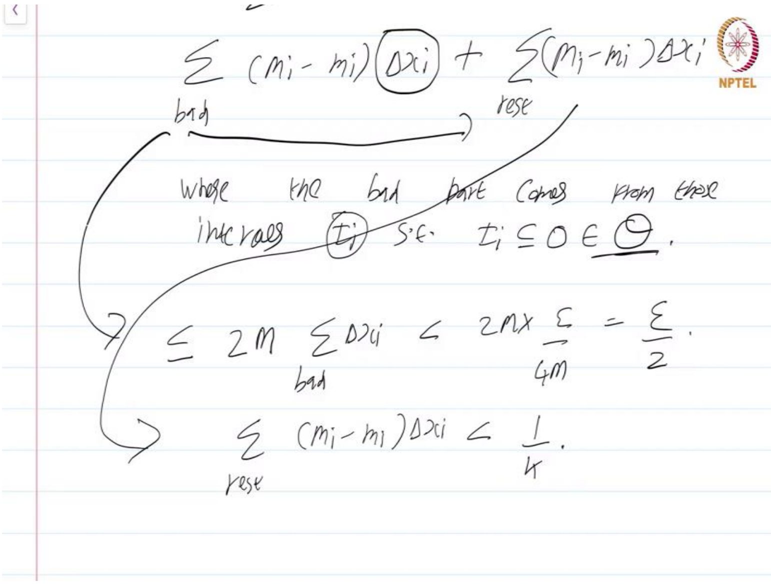


You have this a, b, you have this a, b; you are choosing this partition in such a way that no two points, successive points are greater than or equal to λ distance away. They are all less than λ distance away ok. So, let I_i be an interval determined by the partition by the partition ok.

Since, the net size of this interval is less than λ it is clear that this $I_i \subset \mathcal{O}$\in $\mathcal{J} \cup \mathcal{O}$. This is the key. So, each one of these intervals determined by the partition is going to be a subset of some element.

So, it is going to be a subset of some element of the open cover $\mathcal{J} \cup \mathcal{O}$, ok. Now, let us get back to the business of estimating of estimating $U(f, P - L(f, P)$. So, $U(f, P - L(f, P) = \sum(M_i - m_i)\Delta x_i$.

(Refer Slide Time: 23:15)



So, you break this sum into two pieces. You write this as again $\sum_{bad}(M_i - m_i)\Delta x_i + \sum_{rest}(M_i - m_i)\Delta x_i$. You are summing up over these intervals and I am saying you break it up into two parts.

The bad part and the rest part, where the bad part comes from those intervals I_i such that $I_i \subset O \in \mathcal{O}$. The part that is there coming from cover of D_k. Now, it might happen that some of these intervals I_i are there as a subset of both an element of fancy $\mathcal{O}$ as well as an element of fancy $\mathcal{I}$.

In that case just make a choice arbitrarily; it really does not matter ok. It really does not matter how; just make sure that each interval appears only once. When you are doing the split up, some intervals might belong to both that can happen ok; that can happen because of redundancies that can happen. Those cases just deal with it arbitrarily; it really does not matter ok. Excellent.

Now, how does this help? Well, think about it for a second. We already know that the sum of all the intervals of this script $\mathcal{O}$, this fancy $\mathcal{O}$ is less than $\frac{\varepsilon}{2M}$. Now, you might understand why we put $\frac{\varepsilon}{2M}$. Look at this first term, this first term is certainly going to be less than or equal to $2M \sum_{bad} \Delta x_i$, ok.

Where did I get 2M from? Well, $M_i - m_i$ can at the max be 2M, where M is the supremum of $|f(x)|$, on the whole interval $[a, b]$ right. And, this we already know is less than $2M\frac{\varepsilon}{2M} = \varepsilon$. Actually, I think I made it 4M for this reason; for precisely this reason which is $\frac{\varepsilon}{2}$.

So, we have control the first term, the bad part by controlling the behavior of Δx_i ok. Now, for the rest of the terms, what do we know? We know that $(M_i - m_i)$ is certainly going to be less than; let us see, it is going to be less than $\frac{1}{k}$, it is going to be less than $\frac{1}{k}$.

So, these terms $\sum(M_i - m_i)\Delta x_i$ is certainly going to be less than $\frac{1}{k}$, ok. Actually, it is not just $\frac{1}{k}$, it is $\frac{(b-a)}{k}$. Where does this $(b - a)$ come from? Well, it is the sum of all the Δx_i's that can be at the max $(b - a)$.

(Refer Slide Time: 26:51)

$$\leq 2M \sum_{bad} \Delta x_i < 2M \times \frac{\varepsilon}{4M} = \frac{\varepsilon}{2}.$$

$$\sum_{rest} (M_i - m_i)\Delta x_i < \frac{b-a}{k} < \frac{\varepsilon}{2}.$$

net upshot is

$$U(f, P) - L(f, P) < \varepsilon.$$

Hence f is integrable.

Because, we are probably going to sum up over not all the intervals, some of them would have gone in the bad part ok. But we already made k so large, that $\frac{1}{k} < \frac{\varepsilon}{2(b-a)}$. So, this will be less than $\frac{\varepsilon}{2}$.

So, net up short is net up short is $U(f, p) - L(f, p) < \varepsilon$. Hence, f is Riemann integrable; hence f is Riemann integrable. So, again a challenging proof but there is a nice idea of balance in this proof. Please go through the details in the notes, where I have written out all the details; of course, for the purposes of the lecture I have to just skip some of the details.

So, please go through it again, this is a complicated theorem. In the next module, we are going to see several consequences of this; pretty much all the standard theorems will fall in your lap with no difficulty. This is a course on Real Analysis and you have just watched the module on the Riemann-Lebesgue theorem.