Differentiation: a review

Definition.

Let $X \subset \mathbb{R}, x \in X, f: X \to \mathbb{R}$. We say that f is differentiable at x if there exists a limit

$$f'(x) := \lim_{y \to x} \frac{f(y) - f(x)}{y - x}.$$

Remark.

The definition makes sense only at cluster points of X. At isolated points, every function is differentiable.

Reformulation of the definition.

f is differentiable at x iff $f(y) = f(x) + (y - x)\phi(y)$, where $\phi: X \to \mathbb{R}$ is continuous at x. In this case, $\phi(x) = f'(x)$.

Proof.

Note that for $y \neq x, \phi(y) = \frac{f(y)-f(x)}{y-x}$. Thus ϕ is continuous at x iff there exists

$$\lim_{y \to x} \frac{f(y) - f(x)}{y - x} ■$$

Corollary.

If f is differentiable at x then f is continuous at x.

Proof.

f is a sum of a constant and a multiple of a function continuous at x■

I will not review for the differentiation of the sums, products, and fractions. But the Chain Rule is usually given an incorrect proof in Calculus.

Chain rule.

Let $f: X \to \mathbb{R}, g: Y \to \mathbb{R}$ are two functions, and let $h := g \circ f$. Let f be differentiable at $x \in X$, $f(x) \in Y$, and g be differentiable at $f(x)$. Then h is differentiable at x, and $h'(x) = g'(f(x))f'(x)$.

Proof.

Use reformulated definition of differentiability to write

$$h(y) = g(f(y)) = g(f(x)) + (f(x) - f(y))\psi(f(y)) = g(f(x)) + \psi(f(y))\phi(y)(y - x)$$

Here ϕ and ψ are the corresponding functions for f and g:

$$f(y) = f(x) + (y - x)\phi(y)$$
$$g(z) = g(f(x)) + (z - f(x))\psi(z).$$

The function $\psi(f(y))\phi(y)$ is continuous at x, as a combination of continuous functions. Plugging in $y = x$ we get the expression for the derivative.■

Rolle's Theorem.

Let $f: [a, b] \to \mathbb{R}$ be a continuous function, which is differentiable at every point of (a, b). Assume that $f(a) = f(b)$. Then $\exists c \in (a, b): f'(c) = 0$.

Proof.

Since $[a, b]$ is compact, f reaches its maximum at a point of $[a, b]$. The rest of the proof proceeds as in the Calculus (cf. the textbook also).■

Cauchy Theorem (Generalized Mean Value Theorem).

Let $f, g: [a, b] \to \mathbb{R}$ be two continuous functions, which are differentiable at every point of (a, b).

$$\exists c \in (a, b): f'(c)(g(b) - g(a)) = g'(c)(f(b) - f(a)).$$

Corollary (Mean Value Theorem).

Take $g(t) \equiv t$ in the previous theorem. Then

$$\exists c \in (a, b): f'(c)(b - a) = (f(b) - f(a)).$$

Proof of the Cauchy Theorem.

Consider the function on $[a, b]$:

$$h(x) := (f(x) - f(a))(g(b) - g(a)) - (f(b) - f(a))(g(x) - g(a)).$$

This function satisfies the conditions of the Rolle's Theorem. So $\exists c \in (a, b): h'(c) = 0$.

Thus $f'(c)(g(b) - g(a)) - g'(c)(f(b) - f(a)) = 0$■

Riemann integration

Definition.
A *partition* P of the interval $[a,b]$ is a finite subset of $[a,b]$ containing a and b.
Interpretation.
We can order elements of P:
$a = x_1 < x_2 < \dots < x_n = b$
Mesh of the partition P is defined as $mesh(P) := \max_{1\le j\le n-1}|x_{j+1} - x_j|$.
Definition.
A partition Q is called a *refinement* of partition P if $P \subset Q$. If P and Q are two partitions, a partition $R := P \cup Q$ is their common refinement.
Definition.
A marked partition of $[a,b]$ is a pair (P, Z), where P is a partition, and Z is any set of points containing one point from each interval corresponding to P.
Remark.
We can order elements of Z as $\zeta_1 < \zeta_2 < \dots < \zeta_{n-1}$ such that $\zeta_j \in [x_j x_{j+1}]$.

Definition.
Let $f: [a,b] \to \mathbb{R}$. For a marked partition $(P, \mathrm{Z}) = (\{x_1, x_2, \dots, x_n\}, \{\zeta_1, \zeta_2, \dots, \zeta_n\})$ define the Riemann sum for (P, Z) as
$I(f, P, \mathrm{Z}) := \sum_{j=1}^{n-1} f(\zeta_j)(x_{j+1} - x_j)$.

Assume now that f is bounded on $[a,b]$, and $P = \{a = x_1 < x_2 < \dots < x_n = b\}$ is a partition of [a,b].
Define
$$M_j(f,P) := \sup_{x\in[x_j, x_{j+1}]} f(x),\ m_j(f,P) := \inf_{x\in[x_j, x_{j+1}]} f(x),$$
Definition.
The *upper sum* of f with respect to a partition P is defined as
$$U(f,P) := \sum_{j=1}^{n-1} M_j(f,P)(x_{j+1} - x_j).$$
The *lower sum* of f with respect to a partition P is defined as
$$L(f,P) := \sum_{j=1}^{n-1} m_j(f,P)(x_{j+1} - x_j).$$

Remark.
Note that for any marking Z of a partition P we have
$L(f,P) \le I(f,P,\mathrm{Z}) \le U(f,P)$.
Moreover,
$$U(f,P) = \sup_Z I(f,P,Z),\ \ L(f,P) = \inf_Z I(f,P,Z).$$
Proof is left as an exercise.

Lemma.
Let R be a refinement of partition P. Then
$L(f,P) \le L(f,R) \le U(f,R) \le U(f,P)$
Proof.
It is enough to prove the lemma for the case when $R\backslash \mathrm{P}$ consists of exactly one element, and then use the induction on the number of elements in $R\backslash \mathrm{P}$.
Let $R\backslash \mathrm{P} = \{z\}$, and let $x_{j-1} < z < x_j$. Then
$$U(f,P) - U(f,R) = \Big(\sup_{x\in[x_j, x_{j+1}]} f(x)\Big)(x_{j+1} - x_j) - \Big(\sup_{x\in[x_{j-1}, z]} f(x)\Big)(z - x_j) - \Big(\sup_{x\in[z, x_j]} f(x)\Big)(x_{j+1} - z) \ge 0,$$
since

$(x_{j+1} - x_j) = (x_{j+1} - z) + (z - x_j)$,
and
$\sup_{x\in[x_j,x_{j+1}]} f(x) \geq \sup_{x\in[x_j,z]} f(x)$; $\sup_{x\in[x_j,x_{j+1}]} f(x) \geq \sup_{x\in[z,x_{j+1}]} f(x)$.
The inequality for the lower sums is proven the same way.■

Corollary.

For any two partitions P and Q,
$U(f,P) \geq L(f,Q)$

Proof.

Let $R := P \cup Q$ be the common refinement of both P and Q.
Then, by previous lemma,
$L(f,Q) \leq L(f,R) \leq U(f,R) \leq U(f,P)$■

Definition.

Define $U(f) = \inf_P U(f,P)$, $L(f) = \sup_P L(f,P)$: the upper and lower integral of f over $[a,b]$.

Remark.

By the previous corollary, $U(f) \geq L(f)$.

Definition.

f is called *Riemann integrable on* $[a,b]$ if $U(f) = L(f) =: \int_a^b f(x)dx$.

Riemann condition for integrability.

f is Riemann integrable on $[a,b]$ iff $\forall\varepsilon > 0$ one can find a partition P, such that
$U(f,P) - L(f,P) < \varepsilon$.

Proof.

Since $L(f,P) \leq L(f) \leq U(f) \leq U(f,P)$, we have $U(f) - L(f) \leq U(f,P) - L(f,P)$.
Thus Riemann condition implies that $\forall\varepsilon > 0$ $U(f) - L(f) < \varepsilon$. Thus $U(f) = L(f)$.
If f is integrable, then, by the properties of supremum and infimum,
$\forall\varepsilon > 0$ one can find two partitions Q, R such that
$U(f) \leq U(f,Q) < U(f) + \varepsilon/2, L(f) - \varepsilon/2 < L(f,R) \leq L(f)$.

Take $P = R \cup Q$. By previous lemma,
$L(f) - \varepsilon/2 < L(f,R) \leq L(f,P) \leq U(f,P) \leq U(f,Q) < U(f) + \varepsilon/2$.

Since f is integrable, $U(f) = L(f)$, so $U(f,P) - L(f,P) < \varepsilon$. ■

Definition.

Let I be an interval. *Oscilation* of a function f on I is defined as
$osc_I(f) := \sup_{x,y\in I} |f(x) - f(y)| = \sup_{x\in I} f(x) - \inf_{x\in I} f(x)$.
Note that
$$U(f,P) - L(f,P) = \sum_{j=1}^{n-1} osc_{[x_{j-1},x_j]}(f)(x_j - x_{j-1}).$$
Thus f is integrable on $[a,b]$ iff $\forall\varepsilon > 0$ one can find a partition P, such that
$\sum_{j=1}^{n-1} osc_{[x_{j-1},x_j]}(f)(x_j - x_{j-1}) < \varepsilon$.

Theorem.

f is integrable on $[a,b]$ iff
$\forall\varepsilon > 0 \exists\delta > 0$: for any marked partition (P,Z) with $mesh(P) < \delta$
we have $\left|I(f,P,Z) - \int_a^b f(t)dt\right| < \varepsilon$.

Proof.

Note that since
$U(f,P) = \sup_Z I(f,P,Z)$, $L(f,P) = \inf_Z I(f,P,Z)$,
and for an integrable function f
$$L(f,P) \leq \int_a^b f(t)dt \leq U(f,P),$$
the Theorem is equivalent to the following statement
f is integrable on $[a,b]$ iff
$\forall\varepsilon > 0 \exists\delta > 0$: for any partition P with $mesh(P) < \delta$
we have $U(f,P) - L(f,P) < \varepsilon$.
The "only if" part follows immediately from the Riemann integrability condition.

Assume now that f is integrable. Let $M := \sup_{x\in[a,b]}|f(x)|$, and fix $\varepsilon > 0$.
Then the Riemann integrability condition implies that for some partition
$Q = \{x_1, x_2, \dots, x_n\}$ such that
$U(f,Q) - L(f,Q) < \varepsilon/3$.
Let us now take $\delta := \frac{\varepsilon}{3Mn}$, and consider any partition $P = \{y_1, y_2, \dots, y_m\}$ with $mesh(P) < \delta$.
Let us divide the intervals associated with P into two classes: the intervals which lie inside one of the intervals from partition Q and all other intervals. Note that an interval $[y_j, y_{j+1}]$ is in the second class iff for some x_k we have $y_j < x_k < y_{j+1}$. Thus there are at most n intervals of the second class.
Note that if $[y_j, y_{j+1}] \subset [x_k, x_{k+1}]$, then $M_j(f,P) \le M_k(f,Q)$.
Thus

$$\sum_{Class\ I} M_j(f,P)(y_{j+1} - y_j) \le \sum_{k=1}^{n-1} M_k(f,P)(x_{k+1} - x_k) = U(f,Q).$$

On the other hand,
$\sum_{Class\ II} M_j(f,P)(y_{j+1} - y_j) \le M\delta\#(intervals\ of\ Class\ II) \le M\delta n = \varepsilon/3$.
Thus
U(f,P)=$\sum_{Class\ I} M_j(f,P)(y_{j+1} - y_j) + \sum_{Class\ II} M_j(f,P)(y_{j+1} - y_j) \le U(f,Q) + \varepsilon/3$.
Similarly,

$$L(f,P) \ge L(f,Q) - \frac{\varepsilon}{3}.$$

Thus we get $U(f,P) - L(f,P) < \varepsilon$■

Definition.

Let $X \subset \mathbb{R}, f: X \to \mathbb{R}$. f is called *increasing* on X if $\forall x, y \in X, x < y \Rightarrow f(x) \ge f(y)$.
f is called *decreasing* on X if $\forall x, y \in X, x < y \Rightarrow f(x) \le f(y)$.
f is called *monotone* on X if f is either increasing or decreasing.

Theorem.

Every monotone function on $[a,b]$ is integrable on $[a,b]$.

Proof.

We'll prove the Theorem for an increasing function f, the prove for a decreasing function is the same.
Notice that f is bounded, since for any $x \in X$, we have $f(a) \le f(x) \le f(b)$.
Take $M := f(b) - f(a)$. If $M = 0$, then the function is constant, so it is integrable. Assume therefore that $M > 0$.
We will use the Riemann integrability criterion. Fix $\varepsilon > 0$. Take any partition $P = \{x_1, x_2, \dots, x_n\}$ with $mesh(P) < \frac{\varepsilon}{M}$.
Notice that $M_j(f,P) = f(x_{j+1}), m_j(f,P) = f(x_j)$, and so

$$\sum_{j=1}^{n-1} osc_{[x_{j-1},x_j]}(f)(x_j - x_{j-1}) = \sum_{j=1}^{n-1} \left(f(x_{j+1}) - f(x_j)\right)(x_j - x_{j-1})$$
$$\le mesh(P) \sum_{j=1}^{n-1} \left(f(x_{j+1}) - f(x_j)\right) < \frac{\varepsilon}{M}(f(b) - f(a)) = \varepsilon$$■

Theorem.

Every continuous function on $[a,b]$ is integrable on $[a,b]$.

Proof.

Notice that f is bounded, since it is a continuous function on compact $[a,b]$.
We will use the Riemann integrability criterion. Fix $\varepsilon > 0$. Then, since f is uniformly continuous on compact $[a,b]$, one can find $\delta > 0$ such that if $|x - y| < \delta$, then $|f(x) - f(y)| \le \frac{\varepsilon}{b-a}$. Take any partition $P = \{x_1, x_2, \dots, x_n\}$ with $mesh(P) < \delta$. Note that for any j, $osc_{[x_{j-1},x_j]}(f) < \frac{\varepsilon}{b-a}$.

Thus

$$\sum_{j=1}^{n-1} osc_{[x_{j-1},x_j]}(f)(x_j - x_{j-1}) \le \frac{\varepsilon}{b-a} \sum_{j=1}^{n-1} (x_{j+1} - x_j) < \frac{\varepsilon}{b-a}(b-a) = \varepsilon$$■

Lemma.

If f and g are integrable functions on $[a,b]$, and $f \le g$, then $\int_a^b f(t)dt \le \int_a^b g(t)dt$.

Proof.
Just notice that for any partition P, $L(f,P) \leq L(g,P)$■

Lemma.
If f is integrable on $[a,b]$, then $|f|$ is also integrable on $[a,b]$, and
$$\left|\int_a^b f(t)dt\right| \leq \int_a^b |f(t)|dt.$$
Proof.
For any $x, y \in [a, b]$ we have $\big||f(x)| - |f(y)|\big| \leq |f(x) - f(y)|$. Thus for any interval I, $osc_I(|f|) \leq osc_I(f)$.
Thus, for any partition P,
$U(f,P) - L(f,P) \geq U(|f|,P) - L(|f|,P)$.
So $|f|$ is also integrable.
Notice now that $-|f| \leq f \leq |f|$, so
$$-\int_a^b |f(t)|dt \leq \int_a^b f(t)dt \leq \int_a^b |f(t)|dt.$$ ■

Corollary.
If f is integrable function on $[a,b]$ bounded by M. Then
$$\left|\int_a^b f(t)dt\right| \leq M(b-a).$$

Fundamental Theorem of Calculus

Theorem (Fundamental Theorem of Calculus).
Let f be an integrable function on an interval $[a,b]$, and let

$$F(x) := \int_a^x f(t)dt.$$

Then F is a continuous function. If f is continuous at some $y \in [a,b]$, then F is differentiable at y, and $F'(y) = f(y)$.

Definition.
Let $f: [a,b] \to \mathbb{R}$. A function $F: [a,b] \to \mathbb{R}$ is called *antiderivative* of f is $F'(x) = f(x)$ for any $x \in [a,b]$.

Corollary.
Let $f: [a,b] \to \mathbb{R}$ be a continuous function. Then it has an antiderivative. If G is an antiderivative of f, then

$$G(b) - G(a) = \int_a^b f(t)dt.$$

Proof of Corollary.
The function $F(x) := \int_a^x f(t)dt$ is an antiderivative of f, by the Theorem.
If G is any other derivative, then $(F-G)' \equiv 0$. By Mean Value Theorem, it means that $F - G \equiv const$, so

$$G(b) - G(a) = F(b) - F(a) = \int_a^b f(t)dt - 0 \blacksquare$$

Proof of Theorem.
In this proof, let $[y,x]$ denote the interval $[y,x]$ if $x > y$, and the interval $[x,y]$ if $x < y$. Similarly

$$\int_y^x f(t)dt := \begin{cases} \int_y^x f(t)dt, \text{if } x > y \\ -\int_x^y f(t)dt \text{, if x<y} \end{cases}.$$

Note now that, by Assignment 8,

$$F(x) - F(y) = \int_a^x f(t)dt - \int_a^y f(t)dt = \int_y^x f(t)dt$$

Since f is bounded by some number M, we get that
$|F(x) - F(y)| \leq M|x-y|$,
which implies that F is uniformly continuous.
Assume now that f is continuous at x. Then , as before

$$\frac{F(x)-F(y)}{x-y} = \frac{\int_a^x f(t)dt - \int_a^y f(t)dt}{x-y} = \frac{\int_y^x f(t)dt}{x-y}.$$

Notice that $f(x) = \frac{\int_y^x f(x)dt}{x-y}$.
Subtracting the two last identities, and using additivity of the integral, we get

$$\frac{F(x)-F(y)}{x-y} - f(x) = \frac{\int_y^x f(t)dt}{x-y} - \frac{\int_y^x f(x)dt}{x-y} = \frac{\int_y^x \big(f(t) - f(x)\big)dt}{x-y}.$$

To prove that the above expression tends to 0, let us fix $\varepsilon > 0$, and select $\delta > 0$, such that $|x - t| < \delta \Rightarrow |f(x) - f(t)| < \varepsilon$.
Note now that if $|y - x| < \delta$ then $\forall t \in [y,x], |t - x| < \delta$, so $|f(t) - f(x)| < \varepsilon$.
Thus

$$\left|\int_y^x \big(f(t) - f(x)\big)dt\right| \leq \varepsilon|y-x|.$$

So

$$\frac{F(x)-F(y)}{x-y} - f(x) \leq \varepsilon \blacksquare$$

Lebesgue Integrability Theorem.

Definition.
A set $X \subset \mathbb{R}$ is called *a set of measure zero* if
$\forall \varepsilon > 0 \exists$ a collection of intervals $\{(c_n, d_n)\}$ such that
$X \subset \cup_{n\in\mathbb{N}} (c_n, d_n)$, and $\sum_{n\in\mathbb{N}}(d_n - c_n) < \varepsilon$.
A set $X \subset \mathbb{R}$ is called *a set of content zero* if
$\forall \varepsilon > 0 \exists$ finite collection of intervals $\{(c_n, d_n)\}$ such that
$X \subset \cup_{n=1}^{N}(c_n, d_n)$, and $\sum_{n=1}^{N}(d_n - c_n) < \varepsilon$.
Remark.
Any **compact** set of measure zero has content zero (simply because any infinite open cover has finite subcover). Opposite is always true.

Remark.
Any subset of a set of measure (content) zero has measure (content) zero.
Example.
Any finite set has finite content: if the set consists of N points and $\varepsilon > 0$, cover each point of the set by an interval of the length less than $\frac{\varepsilon}{N}$.
Lemma.
Let (A_n) be a sequence(finite or infinite) of the sets of measure zero. Then $A := \cup A_n$ also has measure zero.
Proof.
Fix $\varepsilon > 0$. Let $\mathcal{F}_n$ denote a collection of intervals of total length at most $\frac{\varepsilon}{2^n}$ which cover A_n. Then the collection $\mathcal{F} := \cup_n \mathcal{F}_n$ covers the set A and has total length at most $\sum_{n=1}^{\infty}\frac{\varepsilon}{2^n} = \varepsilon$■
Examples.
1. $\mathbb{Q}$ has zero measure, but does not have zero content.
2. A set has zero content iff its closure is bounded and have zero measure.
3. *Cantor set* has zero measure (and zero content, since it is compact).

Definition.
A property is said to be valid *almost everywhere(a.e.)* if the set of the points where it is **not** valid has zero measure.

Definition.(Oscillation of a function near a point.)
Let $f: [a, b] \to \mathbb{R}, x \in [a, b]$. *Oscillation of f near x* is defined as
$osc_x f := \inf_{r>0} osc_{(x-r,x+r)} f$.
Remark.
f is continuous at x iff $osc_x f = 0$. You will prove it in the assignment.

Lebesgue Theorem.
A bounded function on $[a, b]$ is Riemann integrable iff it is continuous almost everywhere.
Proof.
Let f be an integrable function. It means, by Riemann Criterium, that for each n one can find a partition P_n such that $U(f, P_n) - L(f, P_n) < 4^{-n}$. Let us call an interval I from P_n *wild* if $osc_I f \geq 2^{-n}$, otherwise an interval will be called *tame*. Now let us observe that ($|I|$ denotes the length of the interval I)

$$2^{-n} \sum_{\text{wild } I \text{ in } P_n} |I| \leq \sum_{\text{wild } I \text{ in } P_n} osc_I f\, |I| \leq U(f, P_n) - L(f, P_n) < 4^{-n}$$

Thus
$\sum_{\text{wild } I \text{ in } P_n} |I| < 2^{-n}$.
Now let us define
$A_k := \{x: \exists \text{ wild } [y, z] \text{ in some } P_n \text{ such that } n \geq k \text{ and } y < x < z\}$.
Let us consider

$$A := \left(\bigcap A_k\right) \cup \bigcup_n P_n.$$
Let us first show that A has measure zero.
Each P_n is finite, so union of the sequence of them has zero measure.
Each A_k is covered by the wild intervals from P_n with $n \geq k$. The total length of all these intervals does not exceed
$$\sum_{n=k}^{\infty} 2^{-n} = 2^{-k+1}.$$
Since the last expression tends to zero when $k \to \infty$, the set $\bigcap A_k$ has measure zero.
Assume that now $x \notin A$. It means, in particular, that for some N the intervals form $P_n, n > N$, containing x, is tame. Also x itself is not a point of any partition. Fix $\varepsilon > 0$, and pick $n > N$ such that $2^{-n} < \varepsilon$. Then x lies inside one of the tame intervals from P_n, i.e. $x \in (x_j, x_{j+1})$. Choose small r, such that $(x - r, x + r) \subset (x_j, x_{j+1})$. Thus we get
$osc_x f \leq osc_{(x-r,x+r)} f \leq osc_{(x_j,x_{j+1})} f < 2^{-n} < \varepsilon$, since
(x_j, x_{j+1}) is tame. Since $osc_x f < \varepsilon$ for every $\varepsilon > 0$, we get that $osc_x f = 0$.
So the function f is continuous at all points outside of A. Since A has measure zero,
f is a.e. continuous.

Assume now that f is bounded and a.e. continuous. Let $|f| \leq M$ for some $M > 0$.
Fix $\varepsilon > 0$. We will find a partition P with $U(f,P) - L(f,P) < \varepsilon$.
Let us consider $A := \left\{x\colon\ osc_x f \geq \frac{\varepsilon}{2(b-a)}\right\}$. f is not continuous at any point of A. Thus A has measure zero.
Let us now show that A is closed. Let x be a limit point of A. Fix $r > 0$. Then
$(x - r, x + r) \cap A \neq \emptyset$. Let $y \in (x - r, x + r) \cap A$.
Then for some $\delta > 0$, $(y - \delta, y + \delta) \subset (x - r, x + r)$.
We have
$$\frac{\varepsilon}{2(b-a)} \leq osc_y f \leq osc_{(y-\delta,y+\delta)} f \leq osc_{(x-r,x+r)} f.$$
This implies that
$$osc_x f = \inf osc_{(x-r,x+r)} f \geq \frac{\varepsilon}{2(b-a)}.$$
So $x \in A$.
We just proved that A contains all of its limit points, and so it is a closed set. It is also bounded, since $A \subset [a,b]$. It means that it is compact.
Since A is a compact set of measure zero, it also has a zero content.
Let now $\{(a_k, b_k)\}$ is a finite covering of A with $\sum_{k=1}^{n}(b_k - a_k) < \varepsilon/4M$.
The set $B := [a,b] \setminus \bigcup_{k=1}^{n}(a_k, b_k)$ is compact, and no point of it belongs to A.
Thus $\forall x \in B \exists r_x\colon osc_{(x-r_x,x+r_x)} f < \frac{\varepsilon}{2(b-a)}$.
The family of open sets $\{(x - r_x/2, x + r_x/2)\}_{x \in B}$ is an open cover of B, and we can select a finite subcover $\{(c_j, d_j)\}_{j=1}^{m}$. Note that by the our construction,
$osc_{[c_j,d_j]} f < \frac{\varepsilon}{2(b-a)}$ (since $[c_j, d_j] \subset (x - r_x, x + r_x)$ for some $x \in B$).
Now let us define the partition
$P := \{a, b\} \cup \{a_k, 1 \leq k \leq n\} \cup \{b_k, 1 \leq k \leq n\} \cup \{c_j, 1 \leq j \leq m\} \cup \{d_j, 1 \leq j \leq m\}$.
Call the interval $[x_i, x_{i+1}]$ of the partition P *nice* if $[x_i, x_{i+1}] \subset B$. Otherwise, let us call *it nasty.*
Notice that any nice interval is subset of some $[c_j, d_j]$, and thus
$osc_{[x_i,x_{i+1}]} f < \frac{\varepsilon}{2(b-a)}$. On the other hand, every nast interval subset of some $[a_k, b_k]$, and thus the total length of nasty intervals does not exceed $\frac{\varepsilon}{4M}$.
Now we are ready to estimate

$$U(f,P) - L(f,P) = \sum_i osc_{[x_{i-1},\, x_i]}(f)(x_i - x_{i-1}) =$$

$$\sum_{nice} osc_{[x_{i-1},\, x_i]}(f)(x_i - x_{i-1}) + \sum_{nasty} osc_{[x_{i-1},\, x_i]}(f)(x_i - x_{i-1}) \le$$

$$\frac{\varepsilon}{2(b-a)} \sum_{nice} (x_i - x_{i-1}) + 2M \sum_{nasty} (x_i - x_{i-1}) \le \frac{\varepsilon}{2(b-a)}(b-a) + 2M\frac{\varepsilon}{4M} = \varepsilon. \blacksquare$$