NOTE ON MATH 2060: MATHEMATICAL ANALYSIS II: 2018-19

CHI-WAI LEUNG

1. Riemann Integrable Functions

We will use the following notation throughout this chapter.

(i): All functions $f, g, h...$ are bounded real valued functions defined on $[a, b]$ and $m \leq f \leq M$ on $[a, b]$.
(ii): Let $P : a = x_0 < x_1 < .... < x_n = b$ denote a partition on $[a, b]$; Put $\Delta x_i = x_i - x_{i-1}$ and $\|P\| = \max \Delta x_i$.
(iii): $M_i(f, P) := \sup\{f(x) : x \in [x_{i-1}, x_i\}$; $m_i(f, P) := \inf\{f(x) : x \in [x_{i-1}, x_i\}$.
Set $\omega_i(f, P) = M_i(f, P) - m_i(f, P)$.
(iv): (the *upper sum* of f): $U(f, P) := \sum M_i(f, P)\Delta x_i$
(the *lower sum* of f). $L(f, P) := \sum m_i(f, P)\Delta x_i$.

Remark 1.1. *It is clear that for any partition on $[a, b]$, we always have*
(i) $m(b - a) \leq L(f, P) \leq U(f, P) \leq M(b - a)$.
(ii) $L(-f, P) = -U(f, P)$ *and* $U(-f, P) = -L(f, P)$.

The following lemma is the critical step in this section.

Lemma 1.2. *Let P and Q be the partitions on $[a, b]$. We have the following assertions.*
(i) If $P \subseteq Q$, then $L(f, P) \leq L(f, Q) \leq U(f, Q) \leq U(f, P)$.
(ii) We always have $L(f, P) \leq U(f, Q)$.

Proof. For Part (i), we first claim that $L(f, P) \leq L(f, Q)$ if $P \subseteq Q$. By using the induction on $l := \#Q - \#P$, it suffices to show that $L(f, P) \leq L(f, Q)$ as $l = 1$. Let $P : a = x_0 < x_1 < \cdots < x_n = b$ and $Q = P \cup \{c\}$. Then $c \in (x_{s-1}, x_s)$ for some s. Notice that we have

$$m_s(f, P) \leq \min\{m_s(f, Q), m_{s+1}(f, Q)\}.$$

So, we have

$$m_s(f, P)(x_s - x_{s-1}) \leq m_s(f, Q)(c - x_{s-1}) + m_{s+1}(f, Q)(x_s - c).$$

This gives the following inequality as desired.

$$L(f, Q) - L(f, P) = m_s(f, Q)(c - x_{s-1}) + m_{s+1}(f, Q)(x_s - c) - m_s(f, P)(x_s - x_{s-1}) \geq 0. \tag{1.1}$$

Now by considering $-f$ in the Inequality 1.1 above, we see that $U(f, Q) \leq U(f, P)$.
For Part (ii), let P and Q be any pair of partitions on $[a, b]$. Notice that $P \cup Q$ is also a partition on $[a, b]$ with $P \subseteq P \cup Q$ and $Q \subseteq P \cup Q$. So, Part (i) implies that

$$L(f, P) \leq L(f, P \cup Q) \leq U(f, P \cup Q) \leq U(f, Q).$$

The proof is complete. □

Date: March 6, 2019.

The following plays an important role in this chapter.

Definition 1.3. *Let f be a bounded function on $[a,b]$. The upper integral (resp. lower integral) of f over $[a,b]$, write $\overline{\int_a^b} f$ (resp. $\underline{\int_a^b} f$), is defined by*

$$\overline{\int_a^b} f = \inf\{U(f,P) : P \textit{ is a partation on } [a,b]\}.$$

(resp.

$$\underline{\int_a^b} f = \sup\{L(f,P) : P \textit{ is a partation on } [a,b]\}.)$$

Notice that the upper integral and lower integral of f must exist by Remark 1.1.

Proposition 1.4. *Let f and g both are bounded functions on $[a,b]$. With the notation as above, we always have*

(i)

$$\underline{\int_a^b} f \leq \overline{\int_a^b} f.$$

(ii) $\underline{\int_a^b}(-f) = -\overline{\int_a^b} f$.

(iii)

$$\underline{\int_a^b} f + \underline{\int_a^b} g \leq \underline{\int_a^b}(f+g) \leq \overline{\int_a^b}(f+g) \leq \overline{\int_a^b} f + \overline{\int_a^b} g.$$

Proof. Part (i) follows from Lemma 1.2 at once.
Part (ii) is clearly obtained by $L(-f,P) = -U(f,P)$.
For proving the inequality $\underline{\int_a^b} f + \underline{\int_a^b} g \leq \underline{\int_a^b}(f+g) \leq$ first. It is clear that we have $L(f,P) + L(g,P) \leq L(f+g,P)$ for all partitions P on $[a,b]$. Now let P_1 and P_2 be any partition on $[a,b]$. Then by Lemma 1.2, we have

$$L(f,P_1) + L(g,P_2) \leq L(f, P_1 \cup P_2) + L(g, P_1 \cup P_2) \leq L(f+g, P_1 \cup P_2) \leq \underline{\int_a^b}(f+g).$$

So, we have

$$\underline{\int_a^b} f + \underline{\int_a^b} g \leq \underline{\int_a^b}(f+g). \tag{1.2}$$

As before, we consider $-f$ and $-g$ in the Inequality 1.2, we get $\overline{\int_a^b}(f+g) \leq \overline{\int_a^b} f + \overline{\int_a^b} g$ as desired. □

The following example shows the strict inequality in Proposition 1.4 (iii) may hold in general.

Example 1.5. *Define a function $f,g : [0,1] \to \mathbb{R}$ by*

$$f(x) = \begin{cases} 1 & \textit{if } \ x \in [0,1] \cap \mathbb{Q}; \\ -1 & \textit{otherwise.} \end{cases}$$

and

$$g(x) = \begin{cases} -1 & \text{if } x \in [0,1] \cap \mathbb{Q}; \\ 1 & \text{otherwise.} \end{cases}$$

Then it is easy to see that $f + g \equiv 0$ *and*

$$\overline{\int_0^1} f = \overline{\int_0^1} g = 1 \quad \text{and} \quad \underline{\int_0^1} f = \underline{\int_0^1} g = -1.$$

So, we have

$$-2 = \underline{\int_a^b} f + \underline{\int_a^b} g < \underline{\int_a^b} (f+g) = 0 = \overline{\int_a^b} (f+g) < \overline{\int_a^b} f + \overline{\int_a^b} g = 2.$$

We can now reaching the main definition in this chapter.

Definition 1.6. *Let f be a bounded function on $[a,b]$. We say that f is Riemann integrable over $[a,b]$ if $\overline{\int_b^a} f = \underline{\int_a^b} f$. In this case, we write $\int_a^b f$ for this common value and it is called the Riemann integral of f over $[a,b]$.*
Also, write $R[a,b]$ for the class of Riemann integrable functions on $[a,b]$.

Proposition 1.7. *With the notation as above, $R[a,b]$ is a vector space over $\mathbb{R}$ and the integral*

$$\int_a^b : f \in R[a,b] \mapsto \int_a^b f \in \mathbb{R}$$

defines a linear functional, that is, $\alpha f + \beta g \in R[a,b]$ and $\int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g$ for all $f, g \in R[a,b]$ and $\alpha, \beta \in \mathbb{R}$.

Proof. Let $f, g \in R[a,b]$ and $\alpha, \beta \in \mathbb{R}$. Notice that if $\alpha \geq 0$, it is clear that $\overline{\int_a^b} \alpha f = \alpha \overline{\int_a^b} f = \alpha \int_a^b f = \alpha \underline{\int_a^b} f = \underline{\int_a^b} \alpha f$. Also, if $\alpha < 0$, we have $\overline{\int_a^b} \alpha f = \alpha \underline{\int_a^b} f = \alpha \int_a^b f = \alpha \overline{\int_a^b} f = \underline{\int_a^b} \alpha f$. Therefore, we have $\int_a^b \alpha f = \alpha \int_a^b f$ for all $\alpha \in \mathbb{R}$. For showing $f + g \in R[a,b]$ and $\int_a^b (f+g) = \int_a^b f + \int_a^b g$, these will follows from Proposition 1.4 (*iii*) at once. The proof is finished. $\square$

The following result is the important characterization of a Riemann integrable function. Before showing this, we will use the following notation in the rest of this chapter.
For a partition $P : a = x_0 < x_1 < \cdots < x_n = b$ and $1 \leq i \leq n$, put

$$\omega_i(f, P) := \sup\{|f(x) - f(x')| : x, x' \in [x_{i-1}, x_i]\}.$$

It is easy to see that $U(f,P) - L(f,P) = \sum_{i=1}^n \omega_i(f,P) \Delta x_i$.

Theorem 1.8. *Let f be a bounded function on $[a,b]$. Then $f \in R[a,b]$ if and only if for all $\varepsilon > 0$, there is a partition $P : a = x_0 < \cdots < x_n = b$ on $[a,b]$ such that*

$$0 \leq U(f,P) - L(f,P) = \sum_{i=1}^n \omega_i(f,P) \Delta x_i < \varepsilon. \tag{1.3}$$

Proof. Suppose that $f \in R[a,b]$. Let $\varepsilon > 0$. Then by the definition of the upper integral and lower integral of f, we can find the partitions P and Q such that $U(f,P) < \overline{\int_a^b} f + \varepsilon$ and $\underline{\int_a^b} f - \varepsilon < L(f,Q)$. By considering the partition $P \cup Q$, we see that

$$\underline{\int_a^b} f - \varepsilon < L(f,Q) \le L(f,P\cup Q) \le U(f,P\cup Q) \le U(f,P) < \overline{\int_a^b} f + \varepsilon.$$

Since $\int_a^b f = \overline{\int_a^b} f = \underline{\int_a^b} f$, we have $0 \le U(f,P\cup Q) - L(f,P\cup Q) < 2\varepsilon$. So, the partition $P \cup Q$ is as desired.
Conversely, let $\varepsilon > 0$, assume that the Inequality 1.3 above holds for some partition P. Notice that we have

$$L(f,P) \le \underline{\int_a^b} f \le \overline{\int_a^b} f \le U(f,P).$$

So, we have $0 \le \overline{\int_a^b} f - \underline{\int_a^b} f < \varepsilon$ for all $\varepsilon > 0$. The proof is finished. $\square$

Remark 1.9. *Theorem 1.8 tells us that a bounded function f is Riemann integrable over $[a,b]$ if and only if the* "size" *of the discontinuous set of f is arbitrary small.*

Example 1.10. *Let $f : [0,1] \to \mathbb{R}$ be the function defined by*

$$f(x) = \begin{cases} \frac{1}{p} & \textit{if } x = \frac{q}{p}, \textit{ where } p,q \textit{ are relatively prime positive integers;} \\ 0 & \textit{otherwise.} \end{cases}$$

Then $f \in R[0,1]$.
(Notice that the set of all discontinuous points of f, say D, is just the set of all $(0,1] \cap \mathbb{Q}$. Since the set $(0,1] \cap \mathbb{Q}$ is countable, we can write $(0,1] \cap \mathbb{Q} = \{z_1, z_2, ...\}$. So, if we let $m(D)$ be the "size$''$ of the set D, then $m(D) = m(\bigcup_{i=1}^\infty \{z_i\}) = \sum_{i=1}^\infty m(\{z_i\}) = 0$, in here, you may think that the size of each set $\{z_i\}$ is 0.)

Proof. Let $\varepsilon > 0$. By Theorem 1.8, it aims to find a partition P on $[0,1]$ such that

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

Notice that for $x \in [0,1]$ such that $f(x) \ge \varepsilon$ if and only if $x = q/p$ for a pair of relatively prime positive integers p, q with $\frac{1}{p} \ge \varepsilon$. Since $1 \le q \le p$, there are only finitely many pairs of relatively prime positive integers p and q such that $f(\frac{q}{p}) \ge \varepsilon$. So, if we let $S := \{x \in [0,1] : f(x) \ge \varepsilon\}$, then S is a finite subset of $[0,1]$. Let L be the number of the elements in S. Then, for any partition $P : a = x_0 < \cdots < x_n = 1$, we have

$$\sum_{i=1}^n \omega_i(f,P)\Delta x_i = \Big(\sum_{i:[x_{i-1},x_i]\cap S = \emptyset} + \sum_{i:[x_{i-1},x_i]\cap S \ne \emptyset} \Big)\, \omega_i(f,P)\Delta x_i.$$

Notice that if $[x_{i-1},x_i] \cap S = \emptyset$, then we have $\omega_i(f,P) \le \varepsilon$ and thus,

$$\sum_{i:[x_{i-1},x_i]\cap S = \emptyset} \omega_i(f,P)\Delta x_i \le \varepsilon \sum_{i:[x_{i-1},x_i]\cap S = \emptyset} \Delta x_i \le \varepsilon(1-0).$$

On the other hand, since there are at most $2L$ sub-intervals $[x_{i-1},x_i]$ such that $[x_{i-1},x_i] \cap S \ne \emptyset$ and $\omega_i(f,P) \le 1$ for all $i = 1, ..., n$, so, we have

$$\sum_{i:[x_{i-1},x_i]\cap S \ne \emptyset} \omega_i(f,P)\Delta x_i \le 1 \cdot \sum_{i:[x_{i-1},x_i]\cap S \ne \emptyset} \Delta x_i \le 2L\|P\|.$$

We can now conclude that for any partition P, we have

$$\sum_{i=1}^{n} \omega_i(f,P)\Delta x_i \leq \varepsilon + 2L\|P\|.$$

So, if we take a partition P with $\|P\| < \varepsilon/(2L)$, then we have $\sum_{i=1}^{n} \omega_i(f,P)\Delta x_i \leq 2\varepsilon$.
The proof is finished. $\square$

Proposition 1.11. *Let f be a function defined on $[a,b]$. If f is either monotone or continuous on $[a,b]$, then $f \in R[a,b]$.*

Proof. We first show the case of f being monotone. We may assume that f is monotone increasing. Notice that for any partition $P : a = x_0 < \cdots < x_n = b$, we have $\omega_i(f,P) = f(x_i) - f(x_{i-1})$. So, if $\|P\| < \varepsilon$, we have

$$\sum_{i=1}^{n} \omega_i(f,P)\Delta x_i = \sum_{i=1}^{n} (f(x_i)-f(x_{i-1}))\Delta x_i < \|P\| \sum_{i=1}^{n} (f(x_i)-f(x_{i-1})) = \|P\|(f(b)-f(a)) < \varepsilon(f(b)-f(a)).$$

Therefore, $f \in R[a,b]$ if f is monotone.
Suppose that f is continuous on $[a,b]$. Then f is uniform continuous on $[a,b]$. Then for any $\varepsilon > 0$, there is $\delta > 0$ such that $|f(x) - f(x')| < \varepsilon$ as $x, x' \in [a,b]$ with $|x - x'| < \delta$. So, if we choose a partition P with $\|P\| < \delta$, then $\omega_i(f,P) < \varepsilon$ for all i. This implies that

$$\sum_{i=1}^{n} \omega_i(f,P)\Delta x_i \leq \varepsilon \sum_{i=1}^{n} \Delta x_i = \varepsilon(b-a).$$

The proof is complete. $\square$

Proposition 1.12. *We have the following assertions.*

(i) If $f, g \in R[a,b]$ with $f \leq g$, then $\int_a^b f \leq \int_a^b g$.

(ii) If $f \in R[a,b]$, then the absolute valued function $|f| \in R[a,b]$. In this case, we have $|\int_a^b f| \leq \int_a^b |f|$.

Proof. For Part (i), it is clear that we have the inequality $U(f,P) \leq U(g,P)$ for any partition P. So, we have $\int_a^b f = \overline{\int_a^b} f \leq \overline{\int_a^b} g = \int_a^b g$.
For Part (ii), the integrability of $|f|$ follows immediately from Theorem 1.8 and the simple inequality $||f|(x') - |f|(x'')| \leq |f(x') - f(x'')|$ for all $x', x'' \in [a,b]$. Thus, we have $U(|f|,P) - L(|f|,P) \leq U(f,P) - L(f,P)$ for any partition P on $[a,b]$.
Finally, since we have $-f \leq |f| \leq f$, by Part (i), we have $|\int_a^b f| \leq \int_a^b |f|$ at once. $\square$

Proposition 1.13. *Let $a < c < b$. We have $f \in R[a,b]$ if and only if the restrictions $f|_{[a,c]} \in R[a,c]$ and $f|_{[c,b]} \in R[c,b]$. In this case we have*

$$\int_a^b f = \int_a^c f + \int_c^b f. \tag{1.4}$$

Proof. Let $f_1 := f|_{[a,c]}$ and $f_2 := f|_{[c,b]}$.
It is clear that we always have

$$U(f_1,P_1) - L(f_1,P_1) + U(f_2,P_2) - L(f_2,P_2) = U(P,f) - L(f,P)$$

for any partition P_1 on $[a,c]$ and P_2 on $[c,b]$ with $P = P_1 \cup P_2$.
From this, we can show the sufficient condition at once.
For showing the necessary condition, since $f \in R[a,b]$, for any $\varepsilon > 0$, there is a partition Q on $[a,b]$

such that $U(f,Q) - L(f,Q) < \varepsilon$ by Theorem 1.8. Notice that there are partitions P_1 and P_2 on $[a,c]$ and $[c,b]$ respectively such that $P := Q \cup \{c\} = P_1 \cup P_2$. Thus, we have

$$U(f_1,P_1) - L(f_1,P_1) + U(f_2,P_2) - L(f_2,P_2) = U(f,P) - L(f,P) \le U(f,Q) - L(f,Q) < \varepsilon.$$

So, we have $f_1 \in R[a,c]$ and $f_2 \in R[c,b]$.
It remains to show the Equation 1.4 above. Notice that for any partition P_1 on $[a,c]$ and P_2 on $[c,b]$, we have

$$L(f_1,P_1) + L(f_2,P_2) = L(f, P_1 \cup P_2) \le \underline{\int_a^b} f = \int_a^b f.$$

So, we have $\int_a^c f + \int_c^b f \le \int_a^b f$. Then the inverse inequality can be obtained at once by considering the function $-f$. Then the resulted is obtained by using Theorem 1.8. $\square$

Proposition 1.14. *Let f and g be Riemann integrable functions defined ion $[a,b]$. Then the pointwise product function $f \cdot g \in R[a,b]$.*

Proof. We first show that the square function f^2 is Riemann integrable. In fact, if we let $M = \sup\{|f(x)| : x \in [a,b]\}$, then we have $\omega_k(f^2,P) \le 2M\omega_k(f,P)$ for any partition $P : a = x_0 < \cdots < a_n = b$ because we always have $|f^2(x) - f^2(x')| \le 2M|f(x) - f(x')|$ for all $x, x' \in [a,b]$. Then by Theorem 1.8, the square function $f^2 \in R[a,b]$.
This, together with the identity $f \cdot g = \frac{1}{2}((f+g)^2 - f^2 - g^2)$. The result follows. $\square$

Remark 1.15. *In the proof of Proposition 1.14, we have shown that if $f \in R[a,b]$, then so is its square function f^2. However, the converse does not hold. For example, if we consider $f(x) = 1$ for $x \in \mathbb{Q} \cap [0,1]$ and $f(x) = -1$ for $x \in \mathbb{Q}^c \cap [0,1]$, then $f \notin R[0,1]$ but $f^2 \equiv 1$ on $[0,1]$.*

Proposition 1.16. (Mean Value Theorem for Integrals)
Let f and g be the functions defined on $[a,b]$. Assume that f is continuous and g is a non-negative Riemann integrable function. Then, there is a point $\xi \in (a,b)$ such that

$$\int_a^b f(x)g(x)dx = f(\xi)\int_a^b g(x)dx. \tag{1.5}$$

Proof. By the continuity of f on $[a,b]$, there exist two points x_1 and x_2 in $[a,b]$ such that

$$f(x_1) = m := \min f(x); \text{ and } f(x_2) = M := \max f(x).$$

We may assume that $a \le x_1 < x_2 \le b$. From this, since $g \le 0$, we have

$$mg(x) \le f(x)g(x) \le Mg(x)$$

for all $x \in [a,b]$. From this and Proposition 1.14 above, we have

$$m\int_a^b g \le \int_a^b fg \le M\int_a^b g.$$

So, if $\int_a^b g = 0$, then the result follows at once.
We may now suppose that $\int_a^b g > 0$. The above inequality shows that

$$m = f(x_1) \le \frac{\int_a^b fg}{\int_a^b g} \le f(x_2) = M.$$

Therefore, there is a point $\xi \in [x_1,x_2] \subseteq [a,b]$ so that the Equation 1.5 holds by using the Intermediate Value Theorem for the function f. Thus, it remains to show that such element ξ can be chosen in (a,b).

Let $a \le x_1 < x_2 \le b$ be as above.
If x_1 and x_2 can be found so that $a < x_1 < x_2 < b$, then the result is proved immediately since $\xi \in [x_1, x_2] \subset (a,b)$ in this case.
Now suppose that x_1 or x_2 does not exist in (a,b), i.e., $m = f(a) < f(x)$ for all $x \in (a,b]$ or $f(x) < f(b) = M$ for all $x \in [a,b)$.
Claim 1: If $f(a) < f(x)$ for all $x \in (a,b]$, then $\int_a^b fg > f(a)\int_a^b g$ and hence, $\xi \in (a, x_2] \subseteq (a,b]$.
For showing **Claim1**, put $h(x) := f(x) - f(a)$ for $x \in [a,b]$. Then h is continuous on $[a,b]$ and $h > 0$ on $(a,b]$. This implies that $\int_c^d h > 0$ for any subinterval $[c,d] \subseteq [a,b]$. (**Why**?)
On the other hand, since $\underline{\int_a^b} g = \int_a^b g > 0$, there is a partition $P : a = x_0 < \cdots < x_n = b$ so that $L(g,P) > 0$. This implies that $m_k(g,P) > 0$ for some sub-interval $[x_{k-1}, x_k]$. Therefore, we have

$$\int_a^b hg \ge \int_{x_{k-1}}^{x_k} hg \ge m_k(g,P)\int_{x_{k-1}}^{x_k} h > 0.$$

Hence, we have $\int_a^b fg > f(a)\int_a^b g$. **Claim 1** follows.
Similarly, one can show that if $f(x) < f(b) = M$ for all $x \in [a,b)$, then we have $\int_a^b fg < f(b)\int_a^b g$.
This, together with **Claim 1** give us that such ξ can be found in (a,b). The proof is finished. ☐

2. Fundamental Theorem of Calculus

Now if $f \in R[a,b]$, then by Proposition 1.13, we can define a function $F : [a,b] \to \mathbb{R}$ by

$$F(c) = \begin{cases} 0 & \text{if } c = a \\ \int_a^c f & \text{if } a < c \le b. \end{cases} \tag{2.1}$$

Theorem 2.1. Fundamental Theorem of Calculus: *With the notation as above, assume that $f \in R[a,b]$, we have the following assertion.*

- *(i) If there is a continuous function H on $[a,b]$ which is differentiable on (a,b) with $H' = f$, then $\int_a^b f = H(b) - H(a)$. In this case, H is called an indefinite integral of f. (**note**: if H_1 and H_2 both are the indefinite integrals of f, then by the Mean Value Theorem, we have $H_2 = H_1 +$ constant).*
- *(ii) The function F defined as in Eq. 2.1 above is continuous on $[a,b]$. Furthermore, if f is continuous on $[a,b]$, then F' exists on (a,b) and $F' = f$ on (a,b).*

Proof. For Part (i), notice that for any partition $P : a = x_0 < \cdots < x_n = b$, then by the Mean Value Theorem, for each $[x_{i-1}, x_i]$, there is $\xi \in (x_{i-1}, x_i)$ such that $F(x_i) - F(x_{i-1}) = F'(\xi)\Delta x_i = f(\xi)\Delta x_i$. So, we have

$$L(f,P) \le \sum f(\xi)\Delta x_i = \sum F(x_i) - F(x_{i-1}) = F(b) - F(a) \le U(f,P)$$

for all partitions P on $[a,b]$. This gives

$$\int_a^b f = \underline{\int_a^b} f \le F(b) - F(a) \le \overline{\int_a^b} f = \int_a^b f$$

as desired.
For showing the continuity of F in Part (ii), let $a < c < x < b$. If $|f| \le M$ on $[a,b]$, then we have $|F(x) - F(c)| = |\int_c^x f| \le M(x-c)$. So, $\lim_{x \to c+} F(x) = F(c)$. Similarly, we also have $\lim_{x\to c-} F(x) = F(c)$. Thus F is continuous on $[a,b]$.
Now assume that f is continuous on $[a,b]$. Notice that for any $t > 0$ with $a < c < c+t < b$, we have

$$\inf_{x \in [c,c+t]} f(x) \le \frac{1}{t}(F(c+t) - F(c)) = \frac{1}{t}\int_c^{c+t} f \le \sup_{x\in[c,c+t]} f(x).$$

Since f is continuous at c, we see that $\lim\limits_{t\to 0+} \frac{1}{t}(F(c+t)-F(c)) = f(c)$. Similarly, we have $\lim\limits_{t\to 0-} \frac{1}{t}(F(c+t) - F(c)) = f(c)$. So, we have $F'(c) = f(c)$ as desired. The proof is finished. $\square$

3. Riemann Sums and Change of variables formula

Definition 3.1. *For each bounded function f on $[a,b]$. Call $R(f,P,\{\xi_i\}) := \sum f(\xi_i)\Delta x_i$, where $\xi_i \in [x_{i-1}, x_i]$, the Riemann sum of f over $[a,b]$.*
We say that the Riemann sum $R(f,P,\{\xi_i\})$ converges to a number A as $\|P\| \to 0$, write $A = \lim\limits_{\|P\|\to 0} R(f,P,\{\xi_i\})$, if for any $\varepsilon > 0$, there is $\delta > 0$ such that

$$|A - R(f,P,\{\xi_i\})| < \varepsilon$$

whenever $\|P\| < \delta$ and for any $\xi_i \in [x_{i-1}, x_i]$.

Proposition 3.2. *Let f be a function defined on $[a,b]$. If the limit $\lim\limits_{\|P\|\to 0} R(f,P,\{\xi_i\}) = A$ exists, then f is automatically bounded.*

Proof. Suppose that f is unbounded. Then by the assumption, there exists a partition $P : a = x_0 < \cdots < x_n = b$ such that $|\sum_{k=1}^n f(\xi_k)\Delta x_k| < 1 + |A|$ for any $\xi_k \in [x_{k-1}, x_k]$. Since f is unbounded, we may assume that f is unbounded on $[a, x_1]$. In particular, we choose $\xi_k = x_k$ for $k = 2, ..., n$. Also, we can choose $\xi_1 \in [a, x_1]$ such that

$$|f(\xi_1)|\Delta x_1 > 1 + |A| + |\sum_{k=2}^n f(x_k)\Delta x_k|.$$

It leads to a contradiction because we have $1 + |A| > |f(\xi_1)|\Delta x_1 - |\sum_{k=2}^n f(x_k)\Delta x_k|$. The proof is finished. $\square$

Lemma 3.3. *$f \in R[a,b]$ if and only if for any $\varepsilon > 0$, there is $\delta > 0$ such that $U(f,P) - L(f,P) < \varepsilon$ whenever $\|P\| < \delta$.*

Proof. The converse follows from Theorem 1.8.
Assume that f is integrable over $[a,b]$. Let $\varepsilon > 0$. Then there is a partition $Q : a = y_0 < ... < y_l = b$ on $[a,b]$ such that $U(f,Q) - L(f,Q) < \varepsilon$. Now take $0 < \delta < \varepsilon/l$. Suppose that $P : a = x_0 < ... < x_n = b$ with $\|P\| < \delta$. Then we have

$$U(f,P) - L(f,P) = I + II$$

where

$$I = \sum_{i:Q\cap[x_{i-1},x_i]=\emptyset} \omega_i(f,P)\Delta x_i;$$

and

$$II = \sum_{i:Q\cap[x_{i-1},x_i]\neq\emptyset} \omega_i(f,P)\Delta x_i$$

Notice that we have

$$I \le U(f,Q) - L(f,Q) < \varepsilon$$

and

$$II \le (M-m) \sum_{i:Q\cap[x_{i-1},x_i]\neq\emptyset} \Delta x_i \le (M-m)\cdot 2l \cdot \frac{\varepsilon}{l} = 2(M-m)\varepsilon.$$

The proof is finished. $\square$

Theorem 3.4. *$f \in R[a,b]$ if and only if the Riemann sum $R(f,P,\{\xi_i\})$ is convergent. In this case, $R(f,P,\{\xi_i\})$ converges to $\int_a^b f(x)dx$ as $\|P\| \to 0$.*

Proof. For the proof $(\Rightarrow)$: we first note that we always have

$$L(f,P) \leq R(f,P,\{\xi_i\}) \leq U(f,P)$$

and

$$L(f,P) \leq \int_a^b f(x)dx \leq U(f,P)$$

for any partition P and $\xi_i \in [x_{i-1}, x_i]$.

Now let $\varepsilon > 0$. Lemma 3.3 gives $\delta > 0$ such that $U(f,P) - L(f,P) < \varepsilon$ as $\|P\| < \delta$. Then we have

$$|\int_a^b f(x)dx - R(f,P,\{\xi_i\})| < \varepsilon$$

as $\|P\| < \delta$ and $\xi_i \in [x_{i-1}, x_i]$. The necessary part is proved and $R(f,P,\{\xi_i\})$ converges to $\int_a^b f(x)dx$.
For $(\Leftarrow)$: assume that there is a number A such that for any $\varepsilon > 0$, there is $\delta > 0$, we have

$$A - \varepsilon < R(f,P,\{\xi_i\}) < A + \varepsilon$$

for any partition P with $\|P\| < \delta$ and $\xi_i \in [x_{i-1}, x_i]$.
Notice that f is automatically bounded in this case by Proposition 3.2.
Now fix a partition P with $\|P\| < \delta$. Then for each $[x_{i-1}, x_i]$, choose $\xi_i \in [x_{i-1}, x_i]$ such that $M_i(f,P) - \varepsilon \leq f(\xi_i)$. This implies that we have

$$U(f,P) - \varepsilon(b-a) \leq R(f,P,\{\xi_i\}) < A + \varepsilon.$$

So we have shown that for any $\varepsilon > 0$, there is a partition $\mathcal{P}$ such that

$$\overline{\int_a^b} f(x)dx \leq U(f,P) \leq A + \varepsilon(1+b-a). \tag{3.1}$$

By considering $-f$, note that the Riemann sum of $-f$ will converge to $-A$. The inequality 3.1 will imply that for any $\varepsilon > 0$, there is a partition P such that

$$A - \varepsilon(1+b-a) \leq \underline{\int_a^b} f(x)dx \leq \overline{\int_a^b} f(x)dx \leq A + \varepsilon(1+b-a).$$

The proof is finished. □

Theorem 3.5. *Let $f \in R[c,d]$ and let $\phi : [a,b] \longrightarrow [c,d]$ be a strictly increasing C^1 function with $f(a) = c$ and $f(b) = d$.*
Then $f \circ \phi \in R[a,b]$, moreover, we have

$$\int_c^d f(x)dx = \int_a^b f(\phi(t))\phi'(t)dt.$$

Proof. Let $A = \int_c^d f(x)dx$. By Theorem 3.4, we need to show that for all $\varepsilon > 0$, there is $\delta > 0$ such that

$$|A - \sum f(\phi(\xi_k))\phi'(\xi_k)\triangle t_k| < \varepsilon$$

for all $\xi_k \in [t_{k-1}, t_k]$ whenever $Q : a = t_0 < ... < t_m = b$ with $\|Q\| < \delta$.
Now let $\varepsilon > 0$. Then by Lemma 3.3 and Theorem 3.4, there is $\delta_1 > 0$ such that

$$|A - \sum f(\eta_k)\triangle x_k| < \varepsilon \tag{3.2}$$

and

$$\sum \omega_k(f,P)\triangle x_k < \varepsilon \tag{3.3}$$

for all $\eta_k \in [x_{k-1},x_k]$ whenever $P: c = x_0 < ... < x_m = d$ with $\|P\| < \delta_1$.
Now put $x = \phi(t)$ for $t \in [a,b]$.
Now since ϕ and ϕ' are continuous on $[a,b]$, there is $\delta > 0$ such that $|\phi(t) - \phi(t')| < \delta_1$ and $|\phi'(t) - \phi'(t')| < \varepsilon$ for all $t, t'\ in [a,b]$ with $|t - t'| < \delta$.
Now let $Q: a = t_0 < ... < t_m = b$ with $\|Q\| < \delta$. If we put $x_k = \phi(t_k)$, then $P: c = x_0 < .... < x_m = d$ is a partition on $[c,d]$ with $\|P\| < \delta_1$ because ϕ is strictly increasing.
Note that the Mean Value Theorem implies that for each $[t_{k-1},t_k]$, there is $\xi_k^* \in (t_{k-1},t_k)$ such that

$$\triangle x_k = \phi(t_k) - \phi(t_{k-1}) = \phi'(\xi_k^*)\Delta t_k.$$

This yields that

$$|\triangle x_k - \phi'(\xi_k)\triangle t_k| < \varepsilon \Delta t_k \tag{3.4}$$

for any $\xi_k \in [t_{k-1},t_k]$ for all $k = 1,...,m$ because of the choice of δ.
Now for any $\xi_k \in [t_{k-1},t_k]$, we have

$$\begin{aligned} |A - \sum f(\phi(\xi_k))\phi'(\xi_k)\triangle t_k| \le & |A - \sum f(\phi(\xi_k^*))\phi'(\xi_k^*)\triangle t_k| \\ & + |\sum f(\phi(\xi_k^*))\phi'(\xi_k^*)\triangle t_k - \sum f(\phi(\xi_k^*))\phi'(\xi_k)\triangle t_k| \\ & + |\sum f(\phi(\xi_k^*))\phi'(\xi_k)\triangle t_k - \sum f(\phi(\xi_k))\phi'(\xi_k)\triangle t_k| \end{aligned} \tag{3.5}$$

Notice that inequality 3.2 implies that

$$|A - \sum f(\phi(\xi_k^*))\phi'(\xi_k^*)\triangle t_k| = |A - \sum f(\phi(\xi_k^*))\triangle x_k| < \varepsilon.$$

Also, since we have $|\phi'(\xi_k^*) - \phi'(\xi_k)| < \varepsilon$ for all $k = 1,..,m$, we have

$$|\sum f(\phi(\xi_k^*))\phi'(\xi_k^*)\triangle t_k - \sum f(\phi(\xi_k^*))\phi'(\xi_k)\triangle t_k| \le M(b-a)\varepsilon$$

where $|f(x)| \le M$ for all $x \in [c,d]$.
On the other hand, by using inequality 3.4 we have

$$|\phi'(\xi_k)\triangle t_k| \le \triangle x_k + \varepsilon \triangle t_k$$

for all k. This, together with inequality 3.3 imply that

$$\begin{aligned} & |\sum f(\phi(\xi_k^*))\phi'(\xi_k)\triangle t_k - \sum f(\phi(\xi_k))\phi'(\xi_k)\triangle t_k| \\ & \le \sum \omega_k(f,P)|\phi'(\xi_k)\triangle t_k| \ (\because \phi(\xi_k^*), \phi(\xi_k) \in [x_{k-1},x_k]) \\ & \le \sum \omega_k(f,P)(\triangle x_k + \varepsilon \triangle t_k) \\ & \le \varepsilon + 2M(b-a)\varepsilon. \end{aligned}$$

Finally by inequality 3.5, we have

$$|A - \sum f(\phi(\xi_k))\phi'(\xi_k)\triangle t_k| \le \varepsilon + M(b-a)\varepsilon + \varepsilon + 2M(b-a)\varepsilon.$$

The proof is finished. □

4. Improper Riemann Integrals

Definition 4.1. *Let $-\infty < a < b < \infty$.*

(i) Let f be a function defined on $[a,\infty)$. Assume that the restriction $f|_{[a,T]}$ is integrable over $[a,T]$ for all $T > a$. Put $\displaystyle\int_a^\infty f := \lim_{T\to\infty}\int_a^T f$ if this limit exists.
Similarly, we can define $\int_{-\infty}^b f$ if f is defined on $(-\infty, b]$.

(ii) If f is defined on $(a,b]$ and $f|_{[c,b]} \in R[c,b]$ for all $a < c < b$. Put $\displaystyle\int_a^b f := \lim_{c\to a+}\int_c^b f$ if it exists.
Similarly, we can define $\int_a^b f$ if f is defined on $[a,b)$.

(iii) As f is defined on $\mathbb{R}$, if $\int_0^\infty f$ and $\int_{-\infty}^0 f$ both exist, then we put $\int_{-\infty}^\infty f = \int_{-\infty}^0 f + \int_0^\infty f$.

In the cases above, we call the resulting limits the improper Riemann integrals of f and say that the integrals are convergent.

Example 4.2. *Define (*formally*) an improper integral $\Gamma(s)$ (called the Γ-function) as follows:*

$$\Gamma(s) := \int_0^\infty x^{s-1}e^{-x}dx$$

for $s \in \mathbb{R}$. Then $\Gamma(s)$ is convergent if and only if $s > 0$.

Proof. Put $I(s) := \int_0^1 x^{s-1}e^{-x}dx$ and $II(s) := \int_1^\infty x^{s-1}e^{-x}dx$. We first claim that the integral $II(s)$ is convergent for all $s \in \mathbb{R}$.
In fact, if we fix $s \in \mathbb{R}$, then we have

$$\lim_{x\to\infty}\frac{x^{s-1}}{e^{x/2}} = 0.$$

So there is $M > 1$ such that $\frac{x^{s-1}}{e^{x/2}} \le 1$ for all $x \ge M$. Thus we have

$$0 \le \int_M^\infty x^{s-1}e^{-x}dx \le \int_M^\infty e^{-x/2}dx < \infty.$$

Therefore we need to show that the integral $I(s)$ is convergent if and only if $s > 0$.
Note that for $0 < \eta < 1$, we have

$$0 \le \int_\eta^1 x^{s-1}e^{-x}dx \le \int_\eta^1 x^{s-1}dx = \begin{cases} \frac{1}{s}(1-\eta^s) & \text{if } s-1 \ne -1; \\ -\ln\eta & \text{otherwise .}\end{cases}$$

Thus the integral $I(s) = \displaystyle\lim_{\eta\to 0+}\int_\eta^1 x^{s-1}e^{-x}dx$ is convergent if $s > 0$.
Conversely, we also have

$$\int_\eta^1 x^{s-1}e^{-x}dx \ge e^{-1}\int_\eta^1 x^{s-1}dx = \begin{cases} \frac{e^{-1}}{s}(1-\eta^s) & \text{if } s-1 \ne -1; \\ -e^{-1}\ln\eta & \text{otherwise .}\end{cases}$$

So if $s \le 0$, then $\int_\eta^1 x^{s-1}e^{-x}dx$ is divergent as $\eta \to 0+$. The result follows. $\square$

References

[1] R.G. Bartle and D.R. Sherbert, Introduction to real analysis, Fourth edition, Wiley, (2011).