NOTE ON MATH 2060: MATHEMATICAL ANALYSIS II: 2016-17

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 $m \le f \le M$ 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) $m(b - a) \le L(f, P) \le U(f, P) \le M(b - a)$.
(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* $L(f, P) \le L(f, Q) \le U(f, Q) \le U(f, P)$.
(ii) We always have $L(f, P) \le U(f, Q)$.

Proof. For Part (i), we first claim that $L(f, P) \le L(f, Q)$ if $P \subseteq Q$. By using the induction on $l := \#Q - \#P$, it suffices to show that $L(f, P) \le L(f, Q)$ 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

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

So, we have

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

This gives the following inequality as desired.

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

Now by considering $-f$ in the Inequality 1.1 above, we see that $U(f, Q) \le U(f, P)$.
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

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

The proof is complete. □

Date: April 9, 2018.

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*

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

(resp.

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

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*

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

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. ☐

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

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

Since $\int_a^b f = \overline{\int_a^b} f = \underline{\int_a^b} f$, we have $0 \leq U(f,P\cup Q) - L(f,P\cup Q) < 2\varepsilon$. 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

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

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

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*

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

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 $f(x) \geq \varepsilon$ if and only if $x = q/p$ for a pair of relatively prime positive integers p, q with $\frac{1}{p} \geq \varepsilon$. Since $1 \leq q \leq p$, there are only finitely many pairs of relatively prime positive integers p and q such that $f(\frac{q}{p}) \geq \varepsilon$. So, if we let $S := \{x \in [0,1] : f(x) \geq \varepsilon\}$, 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

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

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

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

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

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

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. □

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. □

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. □

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

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

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

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

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

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

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

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

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

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

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

as desired.
For showing the continuity of F in Part (ii), let $a < c < x < b$. If $|f| \leq M$ on $[a,b]$, then we have $|F(x) - F(c)| = |\int_c^x f| \leq M(x-c)$. 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

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

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

3. Riemann Sums

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. □

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. □

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

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

and

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

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 $M_i(f,P)-\varepsilon\le f(\xi_i)$. This implies that we have

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

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

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

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

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

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

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

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

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

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

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

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

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

Finally by inequality 3.5, we have

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

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 $\int_a^\infty f := \lim_{T\to\infty} \int_a^T f$ 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 $\frac{x^{s-1}}{e^{x/2}} \leq 1$ for all $x \geq M$. Thus we have

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

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

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

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

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

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

5. Uniform Convergence of a Sequence of Differentiable Functions

Proposition 5.1. *Let $f_n : (a,b) \longrightarrow \mathbb{R}$ be a sequence of functions. Assume that it satisfies the following conditions:*

(i) : $f_n(x)$ point-wise converges to a function $f(x)$ on (a,b);
(ii) : each f_n is a C^1 function on (a,b);
(iii) : $f_n' \to g$ uniformly on (a,b).

Then f is a C^1-function on (a,b) with $f' = g$.

Proof. Fix $c \in (a,b)$. Then for each x with $c < x < b$ (similarly, we can prove it in the same way as $a < x < c$), the Fundamental Theorem of Calculus implies that

$$f_n(x) = \int_c^x f'(t)dt + f_n(c).$$

Since $f_n' \to g$ uniformly on (a, b), we see that

$$\int_c^x f_n'(t)dt \longrightarrow \int_c^x g(t)dt.$$

This gives

$$f(x) = \int_c^x g(t)dt + f(c). \tag{5.1}$$

for all $x \in (c, b)$. Similarly, we have $f(x) = \int_c^x g(t)dt + f(c)$ for all $x \in (a, b)$.
On the other hand, g is continuous on (a, b) since each f_n' is continuous and $f_n' \to g$ uniformly on (a, b). Equation 5.1 will tell us that f' exists and $f' = g$ on (a, b). The proof is finished. □

Proposition 5.2. *Let (f_n) be a sequence of differentiable functions defined on (a, b). Assume that*

(i): there is a point $c \in (a, b)$ such that $\lim f_n(c)$ exists;
(ii): f_n' converges uniformly to a function g on (a, b).

Then

(a): f_n converges uniformly to a function f on (a, b);
(b): f is differentiable on (a, b) and $f' = g$.

Proof. For Part (a), we will make use the Cauchy theorem.
Let $\varepsilon > 0$. Then by the assumptions (i) and (ii), there is a positive integer N such that

$$|f_m(c) - f_n(c)| < \varepsilon \quad \text{and} \quad |f_m'(x) - f_n'(x)| < \varepsilon$$

for all $m, n \geq N$ and for all $x \in (a, b)$. Now fix $c < x < b$ and $m, n \geq N$. To apply the Mean Value Theorem for $f_m - f_n$ on (c, x), then there is a point ξ between c and x such that

$$f_m(x) - f_n(x) = f_m(c) - f_n(c) + (f_m'(\xi) - f_n'(\xi))(x - c). \tag{5.2}$$

This implies that

$$|f_m(x) - f_n(x)| \leq |f_m(c) - f_n(c)| + |f_m'(\xi) - f_n'(\xi)||x - c| < \varepsilon + (b - a)\varepsilon$$

for all $m, n \geq N$ and for all $x \in (c, b)$. Similarly, when $x \in (a, c)$, we also have

$$|f_m(x) - f_n(x)| < \varepsilon + (b - a)\varepsilon.$$

So Part (a) follows.
Let f be the uniform limit of (f_n) on (a, b)
For Part (b), we fix $u \in (a, b)$. We are going to show

$$\lim_{x \to u} \frac{f(x) - f(u)}{x - u} = g(u).$$

Let $\varepsilon > 0$. Since (f_n') is uniformly convergent on (a, b), there is $N \in \mathbb{N}$ such that

$$|f_m'(x) - f_n'(x)| < \varepsilon \tag{5.3}$$

for all $m, n \geq N$ and for all $x \in (a, b)$
Note that for all $m \geq N$ and $x \in (a, b) \setminus \{u\}$, applying the Mean value Theorem for $f_m - f_N$ as before, we have

$$\frac{f_m(x) - f_N(x)}{x - u} = \frac{f_m(u) - f_N(u)}{x - u} + (f_m'(\xi) - f_N'(\xi))$$

for some ξ between u and x.
So Eq.5.3 implies that

$$\left|\frac{f_m(x) - f_m(u)}{x - u} - \frac{f_N(x) - f_N(u)}{x - u}\right| \leq \varepsilon \tag{5.4}$$

for all $m \geq N$ and for all $x \in (a,b)$ with $x \neq u$.
Taking $m \to \infty$ in Eq.5.4, we have

$$|\frac{f(x)-f(u)}{x-u} - \frac{f_N(x)-f_N(u)}{x-u}| \leq \varepsilon.$$

Hence we have

$$\begin{aligned}|\frac{f(x)-f(u)}{x-u} - f_N'(u)| &\leq |\frac{f(x)-f(u)}{x-c} - \frac{f_N(x)-f_N(u)}{x-u}| + |\frac{f_N(x)-f_N(u)}{x-u} - f_N'(u)| \\ &\leq \varepsilon + |\frac{f_N(x)-f_N(u)}{x-u} - f_N'(u)|.\end{aligned}$$

So if we can take $0<\delta$ such that $|\frac{f_N(x)-f_N(u)}{x-u} - f_N'(u)| < \varepsilon$ for $0<|x-u|<\delta$, then we have

$$|\frac{f(x)-f(u)}{x-u} - f_N'(u)| \leq 2\varepsilon \tag{5.5}$$

for $0<|x-u|<\delta$. On the other hand, by the choice of N, we have $|f_m'(y)-f_N'(y)|<\varepsilon$ for all $y\in(a,b)$ and $m\geq N$. So we have $|g(u)-f_N'(u)|\leq\varepsilon$. This together with Eq.5.5 give

$$|\frac{f(x)-f(u)}{x-u} - g(u)| \leq 3\varepsilon$$

as $0<|x-u|<\delta$, that is we have

$$\lim_{x\to u}\frac{f(x)-f(u)}{x-u} = g(u).$$

The proof is finished. □

Remark 5.3. *The uniform convergence assumption of (f_n') in Propositions 5.1 and 5.2 is essential.*

Example 5.4. *Let $f_n(x) := \frac{x}{1+n^2x^2}$ for $x\in(-1,1)$. Then we have*

$$g(x) := \lim_n f_n'(x) := \lim_n \frac{1-n^2x^2}{(1+n^2x^2)^2} = \begin{cases} 0 & \text{if } x\neq 0; \\ 1 & \text{if } x=0. \end{cases}$$

On the other hand, $f_n\to 0$ uniformly on $(-1,1)$. In fact, if $f_n'(1/n)=0$ for all $n=1,2,..,$ then f_n attains the maximal value $f_n(1/n)=\frac{1}{2n}$ at $x=1/n$ for each $n=1,...$ and hence, $f_n\to 0$ uniformly on $(-1,1)$.
So Propositions 5.1 and 5.2 does not hold. Note that (f_n') does not converge uniformly to g on $(-1,1)$.

6. Dini's Theorem

Recall that a subset A of $\mathbb{R}$ is said to be *compact* if for any family open intervals cover $\{J_i\}_{i\in I}$ of A, that is, each J_i is and open interval and $A\subseteq\bigcup_{i\in I}J_i$, we can find finitely many $J_{i_1},...,J_{i_N}$ such that $A\subseteq J_{i_1}\cup\cdots\cup J_{i_N}$.

The following is a very important result.

Theorem 6.1. *A subset A of $\mathbb{R}$ is compact if and only if any sequence (x_n) in A has a convergent subsequence (x_{n_k}) such that $\lim_k x_{n_k}\in A$. In particular, every closed and bounded interval is compact by using the Bolzano-Weierstrass Theorem.*

Proposition 6.2. *(***Dini's Theorem***): Let A be a compact subset of $\mathbb{R}$ and $f_n : A \to \mathbb{R}$ be a sequence of continuous functions defined on A. Suppose that*

(i) for each $x \in A$, we have $f_n(x) \le f_{n+1}(x)$ for all $n = 1, 2...$;
(ii) the pointwise limit $f(x) := \lim_n f_n(x)$ exists for all $x \in A$;
(iii) f is continuous on A.

Then f_n converges to f uniformly on A.

Proof. Let $g_n := f - f_n$ defined on A. Then each g_n is continuous and $g_n(x) \downarrow 0$ pointwise on A. It suffices to show that g_n converges to 0 uniformly on A.
Method I: Suppose not. Then there is $\varepsilon > 0$ such that for all positive integer N, we have

$$g_n(x_n) \ge \varepsilon. \tag{6.1}$$

for some $n \ge N$ and some $x_n \in A$. From this, by passing to a subsequence we may assume that $g_n(x_n) \ge \varepsilon$ for all $n = 1, 2, ...$ Then by using the compactness of A, Theorem 6.1 gives a convergent subsequence (x_{n_k}) of (x_n) in A. Let $z := \lim\limits_k x_{n_k} \in A$. Since $g_{n_k}(z) \downarrow 0$ as $k \to \infty$. So, there is a positive integer K such that $0 \le g_{n_K}(z) < \varepsilon/2$. Since g_{n_K} is continuous at z and $\lim\limits_i x_{n_i} = z$, we have $\lim\limits_i g_{n_K}(x_{n_i}) = g_{n_K}(z)$. So, we can choose i large enough such that $i > K$

$$g_{n_i}(x_{n_i}) \le g_{n_K}(x_{n_i}) < \varepsilon/2$$

because $g_m(x_{n_i}) \downarrow 0$ as $m \to \infty$. This contradicts to the Inequality 6.1.
Method II: Let $\varepsilon > 0$. Fix $x \in A$. Since $g_n(x) \downarrow 0$, there is $N(x) \in \mathbb{N}$ such that $0 \le g_n(x) < \varepsilon$ for all $n \ge N(x)$. Since $g_{N(x)}$ is continuous, there is $\delta(x) > 0$ such that $g_{N(x)}(y) < \varepsilon$ for all $y \in A$ with $|x - y| < \delta(x)$. If we put $J_x := (x - \delta(x), x + \delta(x))$, then $A \subseteq \bigcup_{x \in A} J_x$. Then by the compactness of A, there are finitely many $x_1, ..., x_m$ in A such that $A \subseteq J_{x_1} \cup \cdots \cup J_{x_m}$. Put $N := \max(N(x_1), ..., N(x_m))$. Now if $y \in A$, then $y \in J(x_i)$ for some $1 \le i \le m$. This implies that

$$g_n(y) \le g_{N(x_i)}(y) < \varepsilon$$

for all $n \ge N \ge N(x_i)$. □

References

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

7. Absolutely convergent series

Throughout this section, let (a_n) be a sequence of complex numbers.

Definition 7.1. *We say that a series* $\sum_{n=1}^{\infty} a_n$ *is absolutely convergent if* $\sum_{n=1}^{\infty} |a_n| < \infty$.

Also a convergent series $\sum_{n=1}^{\infty} a_n$ *is said to be conditionally convergent if it is not absolute convergent.*

Example 7.2. Important Example : *The series* $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^\alpha}$ *is conditionally convergent when* $0 < \alpha \leq 1$.
This example shows us that a convergent improper integral may fail to the absolute convergence or square integrable property.
For instance, if we consider the function $f : [1, \infty) \longrightarrow \mathbb{R}$ *given by*

$$f(x) = \frac{(-1)^{n+1}}{n^\alpha} \quad \textit{if} \quad n \leq x < n+1.$$

If $\alpha = 1/2$, *then* $\int_1^\infty f(x)dx$ *is convergent but it is neither absolutely convergent nor square integrable.*

Notation 7.3. *Let* $\sigma : \{1, 2...\} \longrightarrow \{1, 2....\}$ *be a bijection. A formal series* $\sum_{n=1}^{\infty} a_{\sigma(n)}$ *is called an rearrangement of* $\sum_{n=1}^{\infty} a_n$.

Example 7.4. *In this example, we are going to show that there is an rearrangement of the series* $\sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i}$ *is divergent although the original series is convergent. In fact, it is conditionally convergent.*
We first notice that the series $\sum_i \frac{1}{2i-1}$ *diverges to infinity. Thus for each* $M > 0$, *there is a positive integer* N *such that*

$$\sum_{i=1}^{n} \frac{1}{2i-1} \geq M \qquad \cdots\cdots\cdots (*)$$

for all $n \geq N$. *Then there is* $N_1 \in \mathbb{N}$ *such that*

$$\sum_{i=1}^{N_1} \frac{1}{2i-1} - \frac{1}{2} > 1.$$

By using $(*)$ *again, there is a positive integer* N_2 *with* $N_1 < N_2$ *such that*

$$\sum_{i=1}^{N_1} \frac{1}{2i-1} - \frac{1}{2} + \sum_{N_1 < i \leq N_2} \frac{1}{2i-1} - \frac{1}{4} > 2.$$

To repeat the same procedure, we can find a positive integers subsequence (N_k) *such that*

$$\sum_{i=1}^{N_1} \frac{1}{2i-1} - \frac{1}{2} + \sum_{N_1 < i \leq N_2} \frac{1}{2i-1} - \frac{1}{4} + \cdots\cdots\cdots - \sum_{N_{k-1} < i \leq N_k} \frac{1}{2i-1} - \frac{1}{2k} > k$$

for all positive integers k. So if we let $a_n = \frac{(-1)^{n+1}}{n}$*, then one can find a bijection* $\sigma : \mathbb{N} \to \mathbb{N}$ *such that the series* $\sum_{i=1}^{\infty} a_{\sigma(i)}$ *is an rearrangement of the series* $\sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i}$ *and diverges to infinity. The proof is finished.*

Theorem 7.5. *Let* $\sum_{n=1}^{\infty} a_n$ *be an absolutely convergent series. Then for any rearrangement* $\sum_{n=1}^{\infty} a_{\sigma(n)}$ *is also absolutely convergent. Moreover, we have* $\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_{\sigma(n)}$*.*

Proof. Let $\sigma : \{1, 2...\} \longrightarrow \{1, 2...\}$ be a bijection as before.
We first claim that $\sum_n a_{\sigma(n)}$ is also absolutely convergent.
Let $\varepsilon > 0$. Since $\sum_n |a_n| < \infty$, there is a positive integer N such that

$$|a_{N+1}| + \cdots\cdots\cdots + |a_{N+p}| < \varepsilon \qquad \cdots\cdots\cdots (*)$$

for all $p = 1, 2....$ Notice that since σ is a bijection, we can find a positive integer M such that $M > \max\{j : 1 \le \sigma(j) \le N\}$. Then $\sigma(i) \ge N$ if $i \ge M$. This together with $(*)$ imply that if $i \ge M$ and $p \in \mathbb{N}$, we have

$$|a_{\sigma(i+1)}| + \cdots\cdots\cdots |a_{\sigma(i+p)}| < \varepsilon.$$

Thus the series $\sum_n a_{\sigma(n)}$ is absolutely convergent by the Cauchy criteria.
Finally we claim that $\sum_n a_n = \sum_n a_{\sigma(n)}$. Put $l = \sum_n a_n$ and $l' = \sum_n a_{\sigma(n)}$. Now let $\varepsilon > 0$. Then there is $N \in \mathbb{N}$ such that

$$|l - \sum_{n=1}^{N} a_n| < \varepsilon \quad \text{and} \quad |a_{N+1}| + \cdots\cdots + |a_{N+p}| < \varepsilon \cdots\cdots\cdots (**)$$

for all $p \in \mathbb{N}$. Now choose a positive integer M large enough so that $\{1, ..., N\} \subseteq \{\sigma(1), ..., \sigma(M)\}$ and $|l' - \sum_{i=1}^{M} a_{\sigma(i)}| < \varepsilon$. Notice that since we have $\{1, ..., N\} \subseteq \{\sigma(1), ..., \sigma(M)\}$, the condition $(**)$ gives

$$|\sum_{n=1}^{N} a_n - \sum_{i=1}^{M} a_{\sigma(i)}| \le \sum_{N<i<\infty} |a_i| \le \varepsilon.$$

We can now conclude that

$$|l - l'| \le |l - \sum_{n=1}^{N} a_n| + |\sum_{n=1}^{N} a_n - \sum_{i=1}^{M} a_{\sigma(i)}| + |\sum_{i=1}^{M} a_{\sigma(i)} - l'| \le 3\varepsilon.$$

The proof is complete. □

8. Power series

Throughout this section, let

$$f(x) = \sum_{i=0}^{\infty} a_i x^i \qquad \cdots\cdots\cdots\cdots (*)$$

denote a formal power series, where $a_i \in \mathbb{R}$.

Lemma 8.1. *Suppose that there is* $c \in \mathbb{R}$ *with* $c \neq 0$ *such that* $f(c)$ *is convergent. Then*

(i) : $f(x)$ *is absolutely convergent for all* x *with* $|x| < |c|$*.*
(ii) : f *converges uniformly on* $[-\eta, \eta]$ *for any* $0 < \eta < |c|$*.*

Proof. For Part (i), note that since $f(c)$ is convergent, then $\lim a_n c^n = 0$. So there is a positive integer N such that $|a_n c^n| \leq 1$ for all $n \geq N$. Now if we fix $|x| < |c|$, then $|x/c| < 1$. Therefore, we have

$$\sum_{n=1}^{\infty} |a_n||x^n| \leq \sum_{n=1}^{N-1} |a_n||x^n| + \sum_{n\geq N} |a_n c^n||x/c|^n \leq \sum_{n=1}^{N-1} |a_n||x^n| + \sum_{n\geq N} |x/c|^n < \infty.$$

So Part (i) follows.
Now for Part (ii), if we fix $0 < \eta < |c|$,then $|a_n x^n| \leq |a_n \eta|^n$ for all n and for all $x \in [-\eta, \eta]$. On the other hand, we have $\sum_n |a_n \eta^n| < \infty$ by Part (i). So f converges uniformly on $[-\eta, \eta]$ by the M-test. The proof is finished. □

Remark 8.2. *In Lemma 8.9(ii), notice that if $f(c)$ is convergent, it does not imply f converges uniformly on $[-c, c]$ in general.*
For example, $f(x) := 1 + \sum_{n=1}^{\infty} \frac{x^n}{n}$. Then $f(-1)$ is convergent but $f(1)$ is divergent.

Definition 8.3. *Call the set dom $f := \{x \in \mathbb{R} : f(c)$ is convergent $\}$ the domain of convergence of f for convenience. Let $0 \leq r := \sup\{|c| : c \in$ dom $f\} \leq \infty$. Then r is called the radius of convergence of f.*

Remark 8.4. *Notice that by Lemma 8.9, then the domain of convergence of f must be the interval with the end points $\pm r$ if $0 < r < \infty$.*
When $r = 0$, then dom $f = \{0\}$.
Finally, if $r = \infty$, then dom $f = \mathbb{R}$.

Example 8.5. *If $f(x) = \sum_{n=0}^{\infty} n!x^n$, then $r = (0)$. In fact, notice that if we fix a non-zero number x and consider $\lim_n |(n+1)!x^{n+1}|/|n!x^n| = \infty$, then by the ratio test $f(x)$ must be divergent for any $x \neq 0$. So $r = 0$ and dom $f = (0)$.*

Example 8.6. *Let $f(x) = 1 + \sum_{n=1}^{\infty} x^n/n^n$. Notice that we have $\lim_n |x^n/n^n|^{1/n} = 0$ for all x. So the root test implies that $f(x)$ is convergent for all x and then $r = \infty$ and dom $f = \mathbb{R}$.*

Example 8.7. *Let $f(x) = 1 + \sum_{n=1}^{\infty} x^n/n$. Then $\lim_n |x^{n+1}/(n+1)| \cdot |n/x^n| = |x|$ for all $x \neq 0$. So by the ration test, we see that if $|x| < 1$, then $f(x)$ is convergent and if $|x| > 1$, then $f(x)$ is divergent. So $r = 1$. Also, it is known that $f(1)$ is divergent but $f(-1)$ is divergent. Therefore, we have dom $f = [-1, 1)$.*

Example 8.8. *Let $f(x) = \sum x^n/n^2$. Then by using the same argument of Example 8.7, we have $r = 1$. On the other hand, it is known that $f(\pm 1)$ both are convergent. So dom $f = [-1, 1]$.*

Lemma 8.9. *With the notation as above, if $r > 0$, then f converges uniformly on $(-\eta, \eta)$ for any $0 < \eta < r$.*

Proof. It follows from Lemma 8.1 at once. □

Remark 8.10. *Note that the Example 8.7 shows us that f may not converge uniformly on $(-r, r)$. In fact let f be defined as in Example 8.7. Then f does not converges on $(-1, 1)$. In fact, if we let $s_n(x) = \sum_{k=0}^{\infty} a_k x^k$, then for any positive integer n and $0 < x < 1$, we have*

$$|s_{2n}(x) - s_n(x)| = \frac{x^{n+1}}{n+1} + \cdots\cdots + \frac{x^n}{2n}.$$

From this we see that if n is fixed, then $|s_{2n}(x) - s_n(x)| \to 1/2$ as $x \to 1-$. So for each n, we can find $0 < x < 1$ such that $|s_{2n}(x) - s_n(x)| > \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$. Thus f does not converges uniformly on $(-1,1)$ by the Cauchy Theorem.

Proposition 8.11. *With the notation as above, let $\ell = \overline{\lim}\, |a_n|^{1/n}$ or $\lim \dfrac{|a_{n+1}|}{|a_n|}$ provided it exists. Then*

$$r = \begin{cases} \frac{1}{\ell} & \text{if } 0 < \ell < \infty; \\ 0 & \text{if } \ell = \infty; \\ \infty & \text{if } \ell = 0. \end{cases}$$

Proposition 8.12. *With the notation as above if $0 < r \leq \infty$, then $f \in C^\infty(-r,r)$. Moreover, the k-derivatives $f^{(k)}(x) = \sum_{n \geq k} a_k n(n-1)(n-2)\cdot\dots\cdot(n-k+1)x^{n-k}$ for all $x \in (-r,r)$.*

Proof. Fix $c \in (-r,r)$. By Lemma 8.9, one can choose $0 < \eta < r$ such that $c \in (-\eta,\eta)$ and f converges uniformly on $(-\eta,\eta)$.
It needs to show that the k-derivatives $f^{(k)}(c)$ exists for all $k \geq 0$. Consider the case $k = 1$ first.
If we consider the series $\sum_{n=0}^\infty (a_n x^n)' = \sum_{n=1}^\infty n a_n x^{n-1}$, then it also has the same radius r because $\lim_n |na_n|^{1/n} = \lim_n |a_n|^{1/n}$. This implies that the series $\sum_{n=1}^\infty n a_n x^{n-1}$ converges uniformly on $(-\eta,\eta)$. Therefore, the restriction $f|(-\eta,\eta)$ is differentiable. In particular, $f'(c)$ exists and $f'(c) = \sum_{n=1}^\infty n a_n c^{n-1}$.
So the result can be shown inductively on k. □

Proposition 8.13. *With the notation as above, suppose that $r > 0$. Then we have*

$$\int_0^x f(t)dt = \sum_{n=0}^\infty \int_0^x a_n t^n dt = \sum_0^\infty \frac{1}{n+1} a_n x^{n+1}$$

for all $x \in (-r,r)$.

Proof. Fix $0 < x < r$. Then by Lemma 8.9 f converges uniformly on $[0,x]$. Since each term $a_n t^n$ is continuous, the result follows. □

Theorem 8.14. **(Abel)** : *With the notation as above, suppose that $0 < r$ and $f(r)$ (or $f(-r)$) exists. Then f is continuous at $x = r$ (resp. $x = -r$), that is $\lim_{x \to r-} f(x) = f(r)$.*

Proof. Note that by considering $f(-x)$, it suffices to show that the case $x = r$ holds.
Assume $r = 1$.
Notice that if f converges uniformly on $[0,1]$, then f is continuous at $x = 1$ as desired.
Let $\varepsilon > 0$. Since $f(1)$ is convergent, then there is a positive integer such that

$$|a_{n+1} + \cdots\cdots\cdots + a_{n+p}| < \varepsilon$$

for $n \geq N$ and for all $p = 1, 2....$ Note that for $n \geq N$; $p = 1, 2...$ and $x \in [0,1]$, we have

$$\begin{aligned} s_{n+p}(x) - s_n(x) = {} & a_{n+1}x^{n+1} + a_{n+2}x^{n+1} + a_{n+3}x^{n+1} + \cdots\cdots\cdots + a_{n+p}x^{n+1} \\ & + a_{n+2}(x^{n+2} - x^{n+1}) + a_{n+3}(x^{n+2} - x^{n+1}) + \cdots\cdots\cdots + a_{n+p}(x^{n+2} - x^{n+1}) \\ & + a_{n+3}(x^{n+3} - x^{n+2}) + \cdots\cdots\cdots + a_{n+p}(x^{n+3} - x^{n+2}) \\ & \vdots \\ & + a_{n+p}(x^{n+p} - x^{n+p-1}). \end{aligned} \tag{8.1}$$

Since $x \in [0,1]$, $|x^{n+k+1} - x^{n+k}| = x^{n+k} - x^{n+k+1}$. So the Eq.8.1 implies that

$|s_{n+p}(x) - s_n(x)| \leq \varepsilon(x_{n+1} + (x^{n+1} - x^{n+2}) + (x^{n+2} - x^{n+3}) + \cdots + (x^{n+p-1} - x^{n+p})) = \varepsilon(2x^{n+1} - x^{n+p}) \leq 2\varepsilon.$

So f converges uniformly on $[0,1]$ as desired.

Finally for the general case, we consider $g(x) := f(rx) = \sum_n a_n r^n x^n$. Note that $\lim_n |a_n r^n|^{1/n} = 1$ and $g(1) = f(r)$. Then by the case above,, we have shown that

$$f(r) = g(1) = \lim_{x \to 1-} g(x) = \lim_{x \to r-} f(x).$$

The proof is finished. □

Remark 8.15. *In Remark 8.10, we have seen that f may not converges uniformly on $(-r,r)$. However, in the proof of Abel's Theorem above, we have shown that if $f(\pm r)$ both exist, then f converges uniformly on $[-r,r]$ in this case.*

9. Real analytic functions

Proposition 9.1. *Let $f \in C^\infty(a,b)$ and $c \in (a,b)$. Then for any $x \in (a,b) \setminus \{c\}$ and for any $n \in \mathbb{N}$, there is $\xi = \xi(x,n)$ between c and x such that*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k + \int_c^x \frac{f^{(n+1)}(t)}{n!}(x-t)^n dt$$

Call $\sum_{k=0}^{\infty} \frac{f^{(k)}(c)}{k!}(x-c)^k$ (may not be convergent) the Taylor series of f at c.

Proof. It is easy to prove by induction on n and the integration by part. □

Definition 9.2. *A real-valued function f defined on (a,b) is said to be real analytic if for each $c \in (a,b)$, one can find $\delta > 0$ and a power series $\sum_{k=0}^\infty a_k(x-c)^k$ such that*

$$f(x) = \sum_{k=0}^{\infty} a_k(x-c)^k \qquad \cdots\cdots\cdots (*)$$

for all $x \in (c-\delta, c+\delta) \subseteq (a,b)$.

Remark 9.3.

(i) : Concerning about the definition of a real analytic function f, the expression $(*)$ above is uniquely determined by f, that is, each coefficient a_k's is uniquely determined by f. In fact, by Proposition 8.12, we have seen that $f \in C^\infty(a,b)$ and

$$a_k = \frac{f^{(k)}(c)}{k!} \qquad \cdots\cdots\cdots (**)$$

for all $k = 0,1,2,....$

(ii) : Although every real analytic function is C^∞, the following example shows that the converse does not hold.
Define a function $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0; \\ 0 & \text{if } x = 0. \end{cases}$$

One can directly check that $f \in C^\infty(\mathbb{R})$ and $f^{(k)}(0) = 0$ for all $k = 0,1,2....$ So if f is real analytic, then there is $\delta > 0$ such that $a_k = 0$ for all k by the Eq.$(**)$ above and hence $f(x) \equiv 0$ for all $x \in (-\delta, \delta)$. It is absurd.

(iii) **Interesting Fact** : *Let D be an open disc in $\mathbb{C}$. A complex analytic function f on D is similarly defined as in the real case. However, we always have: f is complex analytic if and only if it is C^∞.*

Proposition 9.4. *Suppose that $f(x) := \sum_{k=0}^\infty a_k(x-c)^k$ is convergent on some open interval I centered at c, that is $I = (c-r, c+r)$ for some $r > 0$. Then f is analytic on I.*

Proof. We first note that $f \in C^\infty(I)$. By considering the translation $x - c$, we may assume that $c = 0$. Now fix $z \in I$. Now choose $\delta > 0$ such that $(z-\delta, z+\delta) \subseteq I$. We are going to show that

$$f(x) = \sum_{j=0}^{\infty} \frac{f^{(j)}(z)}{j!}(x-z)^j.$$

for all $x \in (z-\delta, z+\delta)$.
Notice that $f(x)$ is absolutely convergent on I. This implies that

$$\begin{aligned} f(x) &= \sum_{k=0}^{\infty} a_k (x-z+z)^k \\ &= \sum_{k=0}^{\infty} a_k \sum_{j=0}^{k} \frac{k(k-1)\cdot\cdots\cdot(k-j+1)}{j!}(x-z)^j z^{k-j} \\ &= \sum_{j=0}^{\infty} \Big(\sum_{k\geq j} k(k-1)\cdot\cdots\cdot(k-j+1)a_k z^{k-j}\Big) \frac{(x-z)^j}{j!} \\ &= \sum_{j=0}^{\infty} \frac{f^{(j)}(z)}{j!}(x-z)^j \end{aligned}$$

for all $x \in (z-\delta, z+\delta)$. The proof is finished. □

Example 9.5. *Let $\alpha \in \mathbb{R}$. Recall that $(1+x)^\alpha$ is defined by $e^{\alpha \ln(1+x)}$ for $x > -1$.*
Now for each $k \in \mathbb{N}$, put

$$\binom{\alpha}{k} = \begin{cases} \frac{\alpha(\alpha-1)\cdot\cdots\cdot(\alpha-k+1)}{k!} & \text{if } k \neq 0; \\ 1 & \text{if } x = 0. \end{cases}$$

Then

$$f(x) := (1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k$$

whenever $|x| < 1$.
Consequently, $f(x)$ is analytic on $(-1,1)$.

Proof. Notice that $f^{(k)}(x) = \alpha(\alpha-1)\cdot\cdots\cdot(\alpha-k+1)(1+x)^{\alpha-k}$ for $|x| < 1$.
Fix $|x| < 1$. Then by Proposition 9.1, for each positive integer n we have

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k + \int_0^x \frac{f^{(n)}(t)}{(n-1)!}(x-t)^{n-1} dt$$

So by the mean value theorem for integrals, for each positive integer n, there is ξ_n between 0 and x such that

$$\int_0^x \frac{f^{(n)}(t)}{(n-1)!}(x-t)^{n-1} dt = \frac{f^{(n)}(\xi_n)}{(n-1)!}(x-\xi_n)^{n-1} x$$

Now write $\xi_n = \eta_n x$ for some $0 < \eta_n < 1$ and $R_n(x) := \dfrac{f^{(n)}(\xi_n)}{(n-1)!}(x-\xi_n)^{n-1} x$. Then

$$R_n(x) = (\alpha-n+1)\binom{\alpha}{n-1}(1+\eta_n x)^{\alpha-n}(x-\eta_n x)^{n-1} x = (\alpha-n+1)\binom{\alpha}{n-1} x^n (1+\eta_n x)^{\alpha-1}\Big(\frac{1-\eta_n}{1+\eta_n x}\Big)^{n-1}.$$

We need to show that $R_n(x) \to 0$ as $n \to \infty$, that is the Taylor series of f centered at 0 converges to f. By the Ratio Test, it is easy to see that the series $\sum_{k=0}^{\infty} (\alpha-k+1)\binom{\alpha}{k} y^k$ is convergent as $|y| < 1$.

This tells us that $\lim_n |(\alpha-n+1)\binom{\alpha}{n} x^n| = 0$.
On the other hand, note that we always have $0 < 1-\eta_n < 1+\eta_n x$ for all n because $x > -1$. Thus, we

can now conclude that $R_n(x) \to 0$ as $|x| < 1$. The proof is finished. Finally the last assertion follows from Proposition 9.4 at once. The proof is complete. □

References

(Chi-Wai Leung) Department of Mathematics, The Chinese University of Hong Kong, Shatin, Hong Kong
E-mail address: cwleung@math.cuhk.edu.hk