MATH2068 MATHEMATICAL ANALYSIS II (2023-24)

CHI-WAI LEUNG

1. Differentiation

Throughout this section, let I be an open interval (not necessarily bounded) and let f be a real-valued function defined on I.

Definition 1.1. *Let $c \in I$. We say that f is differentiable at c if the following limit exists:*

$$\lim_{x\to c} \frac{f(x)-f(c)}{x-c}.$$

In this case, we write $f'(c)$ for the above limit and we call it the derivative of f at c. We say that if f is differentiable on I if $f'(x)$ exists for every point x in I.

Proposition 1.2. *Let $c \in I$. Then $f'(c)$ exists if and only if there is a function φ defined on I such that the function φ is continuous at c and*

$$f(x) - f(c) = \varphi(x)(x-c)$$

for all $x \in I$.
In this case, $\varphi(c) = f'(c)$.

Proof. Assume that $f'(c)$ exists. Define a function $\varphi : I \to \mathbb{R}$ by

$$\varphi(x) = \begin{cases} \frac{f(x)-f(c)}{x-c} & \text{if } \; x \neq c; \\ f'(c) & \text{if } \; x = c. \end{cases}$$

Clearly, we have $f(x) - f(c) = \varphi(x)(x-c)$ for all $x \in I$. We want to show that the function φ is continuous at c. In fact, let $\varepsilon > 0$, by the definition of the limit of a function, there is $\delta > 0$ such that

$$|f'(c) - \frac{f(x)-f(c)}{x-c}| < \varepsilon$$

whenever $x \in I$ with $0 < |x-c| < \delta$. Therefore, we have $|f'(c)-\varphi(x)| < \varepsilon$ as $x \in I$ with $0 < |x-c| < \delta$. Since $\varphi(c) = f'(c)$, we have $|f'(c) - \varphi(x)| < \varepsilon$ as $x \in I$ with $|x-c| < \delta$, hence the function φ is continuous at c as desired.
The converse is clear since $\varphi(x) = \frac{f(x)-f(c)}{x-c}$ if $x \neq c$. The proof is complete. □

Proposition 1.3. *Using the notation as above, if f is differentiable at c, then f is continuous at c.*

Proof. By using Proposition 1.2, if $f'(c)$ exists, then there is a function φ defined on I such that the function φ is continuous at c and we have $f(x) - f(c) = \varphi(x)(x-c)$ for all $x \in I$. This implies that $\lim_{x\to c} f(x) = f(c)$, so f is continuous at c as desired. □

Remark 1.4. *In general, the converse of Proposition 1.3 does not hold, for example, the function $f(x) := |x|$ is a continuous function on $\mathbb{R}$ but $f'(0)$ does not exist.*

Date: April 19, 2024.

Proposition 1.5. *Let f and g be the functions defined on I. Assume that f and g both are differentiable at $c \in I$. We have the following assertions.*

(i) $(f+g)'(c)$ *exists and* $(f+g)'(c) = f'(c) + g'(c)$.

(ii) The product $(f \cdot g)'(c)$ *exists and* $(f \cdot g)'(c) = f'(c)g(c) + f(c)g'(c)$.

(iii) If $g(c) \neq 0$, *then we have* $(\frac{f}{g})'(c)$ *exists and* $(\frac{f}{g})'(c) = \frac{f'(c)g(c)-f(c)g'(c)}{g(c)^2}$.

Proof. Part (i) clearly follows from the definition of the limit of a function.
For showing Part (ii), note that we have

$$\frac{f(x)g(x) - f(c)g(c)}{x-c} = \frac{f(x)-f(c)}{x-c}g(x) + f(c)\frac{g(x)-g(c)}{x-c}$$

for all $x \in I$ with $x \neq c$. From this, together with Proposition 1.3, Part (ii) follows.
For Part (iii), by using Part (ii), it suffices to show that $(\frac{1}{g})'(c) = -\frac{g'(c)}{g(c)^2}$. In fact, $g'(c)$ exists, so g is continuous at c. Since $g(c) \neq 0$, there is $\delta_1 > 0$ so that $g(x) \neq 0$ for all $x \in I$ with $|x-c| < \delta_1$. Then we have

$$\frac{1}{x-c}(\frac{1}{g(x)} - \frac{1}{g(c)}) = \frac{1}{x-c}(\frac{g(c)-g(x)}{g(x)g(c)})$$

for all $x \in I$ with $0 < |x-c| < \delta_1$. By taking $x \to c$, we see that $(\frac{1}{g})'(c)$ exists and $(\frac{1}{g})'(c) = \frac{-g'(c)}{g(c)^2}$.
The proof is complete. □

Proposition 1.6. (Chain Rule): *Let f, g be functions defined on $\mathbb{R}$. Let $d = f(c)$ for some $c \in \mathbb{R}$. Suppose that $f'(c)$ and $g'(d)$ exist. Then the derivative of composition $(g \circ f)'(c)$ exists and $(g \circ f)'(c) = g'(d)f'(c)$.*

Proof. By using Proposition 1.2, we want to find a function $\varphi : \mathbb{R} \to \mathbb{R}$ such that

$$g \circ f(x) - g \circ f(c) = \varphi(x)(x-c)$$

for all $x \in \mathbb{R}$ and the function $\varphi(x)$ is continuous at c, and so $(g \circ f)'(c) = \varphi(c)$.
Let $y = f(x)$. By using Proposition 1.2 again, there is a function and $\beta(y)$ so that $g(y) - g(d) = \beta(y)(y-d)$ for all $y \in \mathbb{R}$ and $\beta(y)$ is continuous at d. Similarly, there is a function $\alpha(x)$ we have $f(x) - f(c) = \alpha(x)(x-c)$ for all $x \in \mathbb{R}$ and $\alpha(x)$ is continuous at c. These two equations imply that

$$g \circ f(x) - g \circ f(c) = \beta(f(x))(f(x) - f(c)) = \beta(f(x))\alpha(x)(x-c)$$

for all $x \in \mathbb{R}$. Let $\varphi(x) := \beta(f(x)) \cdot \alpha(x)$ for $x \in \mathbb{R}$. Since $\beta(d) = g'(d)$ and $\alpha(c) = f'(c)$, we see that $\varphi(c) = \beta(f(c))\alpha(c) = g'(d)f'(c)$. It remains to show that the function φ is continuous at c. In fact, $f'(c)$ exists, so f is continuous at c, and hence the composition $\beta \circ f(x)$ is continuous at c. In addition, the function α is continuous at c. Therefore, the function $\varphi := (\beta \circ f) \cdot \alpha$ is continuous at c, and so $(g \circ f)'(c)$ exists with $(g \circ f)'(c) = \varphi(c) = g'(d)f'(c)$. The proof is complete. □

Proposition 1.7. *Let I and J be open intervals. Let f be a strictly increasing function from I onto J. Let $d = f(c)$ for $c \in I$. Assume that $f'(c)$ exists and the inverse of f, write $g := f^{-1}$, is continuous at d. If $f'(c) \neq 0$, then $g'(d)$ exists and $g'(d) = \frac{1}{f'(c)}$.*

Proof. Let $y = f(x)$. Note that by using Proposition 1.2, there is a function F on I such that $f(x) - f(c) = F(x)(x-c)$ for all $x \in I$ and F is continuous at c with $F(c) = f'(c) \neq 0$. F is continuous at c, so there are open intervals I_1 and J_1 such that $c \in I_1 \subseteq I$ and $d \in f(I_1) = J_1$, moreover, $F(x) \neq 0$ for all $x \in I_1$. Note that since $f(x) - f(c) = F(x)(x-c)$, we have $y - d = f(g(y)) - f(g(c)) = F(g(y))(g(y) - g(d))$ for all $y \in J_1$. Since $F(x) \neq 0$ for all $x \in I_1$, we have $g(y) - g(d) = F(g(y))^{-1}(y-d)$ for all $y \in J_1$. Note that the function $F(g(y))^{-1}$ is continuous at d. Thus, $g'(d)$ exists and $g'(d) = F(g(d))^{-1} = \frac{1}{f'(c)}$ as desired. □

Definition 1.8. *Let D be a non-empty subset of $\mathbb{R}$ and let g be a real-valued function defined on D.*

(i) We say that g has an absolute maximum (resp. absolute minimum) at a point $c \in D$ if $g(c) \geq g(x)$ (resp. $g(c) \leq g(x)$) for all $x \in D$.
In this case, c is called an absolute extreme point of g.

(ii) We say that g has a local maximum (resp. local minimum) at a point $c \in D$ if there is $r > 0$ such that $(c-r, c+r) \subseteq D$ and $g(c) \geq g(x)$ (resp. $g(c) \leq g(x)$) for all $x \in (c-r, c+r)$.
In this case, c is called a local extreme point of g.

Remark 1.9. *Note that an absolute extreme point of a function g need not be a local extreme point, for example if $g(x) := x$ for $x \in [0,1]$, then g has an absolute maximum point at $x = 1$ of g but 1 is not a local maximum point of g.*

Proposition 1.10. *Let I be an open interval and let f be a function on I. Assume that f has a local extreme point at $c \in I$ and $f'(c)$ exists. Then $f'(c) = 0$.*

Proof. Without lost the generality, we may assume that f has local minimum at c. Then there is $r > 0$ such that $f(x) \geq f(c)$ for $x \in (c-r, c+r) \subseteq I$. Since $f'(c)$ exists, by using Proposition 1.2, there is a function φ defined on I such that $f(x) - f(c) = \varphi(x)(x-c)$ for all $x \in I$ and φ is continuous at c with $\varphi(c) = f'(c)$. Thus, we have $\varphi(x)(x-c) \geq 0$ for all $x \in (c-r, c+r)$. From this we see that $\varphi(x) \geq 0$ as $x \in (c, c+r)$, similarly, $\varphi(x) \leq 0$ as $x \in (c-r, c)$. The function φ is continuous at c, so $\varphi(c) = 0$ and hence $f'(c) = \varphi(c) = 0$ as desired. ☐

Proposition 1.11. Rolle's Theorem*: Let $f : [a,b] \to \mathbb{R}$ be a continuous function. Assume that $f'(x)$ exists for all $x \in (a,b)$ and $f(a) = f(b)$. Then there is a point $c \in (a,b)$ such that $f'(c) = 0$.*

Proof. Recall a fact that every continuous function defined a compact attains absolute points, that is, there are c_1 and c_2 such that $f(c_1) = \min_{x \in [a,b]} f(x)$ and $f(c_2) = \max_{x \in [a,b]} f(x)$, hence, $f(c_1) \leq f(x) \leq f(c_2)$ for all $x \in [a,b]$. If $f(c_1) = f(c_2)$, then $f(x) \equiv f(c_1) = f(c_2)$ for all $x \in [a,b]$, so $f'(x) \equiv 0$ for all $x \in (a,b)$.
Otherwise, suppose that $f(c_1) < f(c_2)$. Since $f(a) = f(b)$, we have $c_1 \in (a,b)$ or $c_2 \in (a,b)$. We may assume that $c_1 \in (a,b)$. Then $x = c_1$ is a local minimum point of f. Therefore, $f'(c_1) = 0$ by using Proposition 1.10. ☐

Theorem 1.12. Main Value Theorem: *If $f : [a,b] \to \mathbb{R}$ is a continuous function and is differentiable on (a,b), then there is a point $c \in (a,b)$ such that $f(b) - f(a) = f'(c)(b-a)$.*

Proof. Define a function $\varphi : [a,b] \to \mathbb{R}$ by

$$\varphi(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b-a}(x-a)$$

for $x \in [a,b]$. Note that the function φ is continuous on $[a,b]$ with $\varphi(a) = \varphi(b) = 0$, in addition, $\varphi'(x)$ exists for all $x \in (a,b)$. The Rolle's Theorem implies that there is a point $c \in (a,b)$ such that

$$0 = \varphi'(c) = f'(c) - \frac{f(b) - f(a)}{b-a}.$$

The proof is complete. ☐

Corollary 1.13. *Assume that $f : [a,b] \to \mathbb{R}$ is a continuous function and is differentiable on (a,b). If $f' \equiv 0$ on (a,b), then f is a constant function.*

Proof. Fix any point $z \in (a,b)$. Let $x \in (z,b]$. By using the Mean Value Theorem, there is a point $c \in (z,x)$ such that $f(x) - f(z) = f'(c)(x-z)$. If $f' \equiv 0$ on (a,b), so $f(x) = f(z)$ for all $x \in [z,b]$. Similarly, we have $f(x) = f(z)$ for all $x \in [a,z]$. The proof is complete. □

Definition 1.14. *We call a function f is a C^1-function on I if $f'(x)$ exists and continuous on I. In addition, we define the n-derivatives of f by $f^{(n)}(x) := f^{(n-1)}(x)$ for $n \geq 2$, provided it exists. In this case, we say that f is a C^n-function on I. In particular, we call f a C^∞-function (or smooth function) if f is a C^n-function for all $n = 1, 2....$*
For example, the exponential function $\exp x$ is a very important example of smooth function on $\mathbb{R}$.

Corollary 1.15. Inverse Mapping Theorem*: Let f be a C^1-function on an open interval I and let $c \in I$. Assume that $f'(c) \neq 0$. Then there is $r > 0$ such that the function f is a strictly monotone function on $(c-r, c+r) \subseteq I$. If we let $J := f(c-r, c+r))$, then the inverse function $g := f^{-1} : J \to (c-r, c+r)$ is also a C^1-function.*

Proof. We may assume that $f'(c) > 0$. $f'(x)$ is continuous on I, so there is $r > 0$ such that $f'(x) > 0$ for all $x \in (c-r, c+r) \subseteq I$. For any x_1 and x_2 in $(c-r,, c+r)$ with $x_1 < x_2$, by using the Mean Value Theorem, we have $f(x_2) - f(x_1) = f'(v)(x_2 - x_1)$ for some $v \in (x_1, x_2)$, and hence $f(x_2) > f(x_1)$. Therefore the restriction of f on $(c-r, c+r)$ is a strictly increasing function, thus, it is an injection. Let $J := f((c-r, c+r))$. Then J is an interval by the Immediate Value Theorem. Moreover, J is an open interval because f is strictly increasing. Also, if we let $g = f^{-1}$ on J, then g is continuous on J due to the fact that every continuous bijection on a compact set is a homeomorphism. Therefore, by Proposition 1.7, we see that $g'(y)$ exists on J and $g'(y) = \frac{1}{f'(x)}$ for $y = f(x)$ and $x \in (c-r, c+r)$. Therefore, g is a C^1 function on J. The proof is complete. □

Proposition 1.16. Cauchy Mean Value Theorem*: Let $f, g : [a,b] \to \mathbb{R}$ be continuous functions with $g(a) \neq g(b)$. Assume that f, g are differentiable functions on (a,b) and $g'(x) \neq 0$ for all $x \in (a,b)$. Then there is a point $c \in (a,b)$ such that $\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)}$.*

Proof. Define a function ψ on $[a,b]$ by $\psi(x) = f(x) - f(a) - \frac{f(b)-f(a)}{g(b)-g(a)}(g(x) - g(a))$ for $x \in [a,b]$. Then by using the similar argument as in the Mean Value Theorem, the result follows. □

Theorem 1.17. Lagrange Remainder Theorem*: Let f be a $C^{(n+1)}$ function defined on (a,b). Let $x_0 \in (a,b)$. Then for each $x \in (a,b)$, there is a point c between x_0 and x such that*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1}.$$

Proof. We may assume that $x_0 < x < b$. **Case:** We first assume that $f^{(k)}(x_0) = 0$ for all $k = 0, 1, ..., n$. Put $g(t) = (t-x_0)^{n+1}$ for $t \in [x_0, x]$. Then $g'(t) = (n+1)(t-x_0)^n$ and $g(x_0) = 0$. Then by the Cauchy Mean Value Theorem, there is $x_1 \in (x_0, x)$ such that $\frac{f(x)}{g(x)} = \frac{f(x)-f(x_0)}{g(x)-g(x_0)} = \frac{f'(x_1)}{g'(x_1)}$. Using the same step for f' and g' on $[x_0, x_1]$, there is $x_2 \in (x_0, x_1)$ such that $\frac{f'(x_1)}{g'(x_1)} = \frac{f'(x_1)-f'(x_0)}{g'(x_1)-g'(x_0)} = \frac{f^{(2)}(x_2)}{g(2)(x_2)}$. To repeat the same step, there are $x_1, x_2, ..., x_{n+1}$ in (a,b) such that $x_k \in (x_0, x_{k-1})$ for $k = 1, 2, ..., n+1$ and

$$\frac{f(x)}{g(x)} = \frac{f'(x_1)}{g'(x_1)} = \cdots = \frac{f^{(n+1)}(x_{n+1})}{g^{(n+1)}(x_{n+1})}.$$

In addition, note that $g^{n+1}(x_{n+1}) = (n+1)!$. Therefore, we have $\frac{f(x)}{g(x)} = \frac{f^{(n+1)}(x_{n+1})}{(n+1)!}$, and hence $f(x) = \frac{f^{(n+1)}(x_{n+1})}{(n+1)!}(x-x_0)^{n+1}$. Note $x_{n+1} \in (x_0, x)$ and thus, the result holds for this case.

For the general case, put $G(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$ for $x \in (a,b)$. Note that we have $G(x_0) = G'(x_0) = \cdots = G^{(n)}(x_0) = 0$. Then by the Claim above, there is a point $c \in (x_0, x)$ such that $G(x) = \frac{G^{(n+1)}(c)}{(n+1)!}$. Since $G^{(n+1)}(c) = f^{(n+1)}(c)$, $f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{f^{(n+1)}(c)}{(n+1)!}$. The proof is complete. □

Example 1.18. Recall that the exponential function e^x is defined by

$$e^x := \sum_{k=0}^{\infty} \frac{x^k}{k!} := \lim_{n\to\infty} \sum_{k=0}^{n} \frac{x^k}{k!}$$

for $x \in \mathbb{R}$. Note that the above limit always exists for all $x \in \mathbb{R}$ (shown in the last chapter).
Show that the natural base e is an irrational number.
Put $f(x) := e^x$ for $x \in \mathbb{R}$. It is a known fact f is a C^∞ function and $f^{(n)}(x) = e^x$ for all $x \in \mathbb{R}$. Fix any $x > 0$. Then by the Lagrange Theorem, for each positive integer n, there is $c_n \in (0, x)$ such that

$$f(x) = \sum_{k=0}^{n} \frac{x^k}{k!} + \frac{e^{c_n}}{(n+1)!} x^{n+1}.$$

In particular, taking $x = 1$, we have

$$0 < \frac{e^{c_n}}{(n+1)!} = e - \sum_{k=0}^{n} \frac{1}{k!} < \frac{3}{(n+1)!}$$

for all positive integer n. Now if $e = p/q$ for some positive integers p and q, and thus, we have

$$0 < \frac{p}{q} - \sum_{k=0}^{n} \frac{1}{k!} < \frac{3}{(n+1)!}$$

for all $n = 1, 2...$ Now we can choose n large enough such that $(n!)\frac{p}{q} \in \mathbb{N}$. It leads to a contradiction because we have

$$0 < (n!)\frac{p}{q} - (n!)\sum_{k=0}^{n} \frac{1}{k!} < \frac{3(n!)}{(n+1)!} = \frac{3}{n+1} < 1.$$

Therefore, e is irrational.

Proposition 1.19. *Let f be a C^2 function on an open interval I and $x_0 \in I$. Assume that $f'(x_0) = 0$. Then f has local maximum (resp. local minimum) at x_0 if $f^{(2)}(x_0) < 0$ (resp. $f^{(2)}(x_0) > 0$).*

Proof. We assume that $f^{(2)}(x_0) > 0$. We want to show that x_0 is a local minimum point of f. The proof of another case is similar. Note that for any $x \in I \setminus \{x_0\}$. Then by the Lagrange Theorem, there is a point c between x_0 and x such that

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2} f^{(2)}(x_0)(x-x_0)^2 = f(x_0) + \frac{1}{2} f^{(2)}(x_0)(x-x_0)^2.$$

$f^{(2)}$ is continuous at x_0 and $f^{(2)}(x_0) > 0$, and so there is $r > 0$ such that $f^{(2)}(x) > 0$ for all $x \in (x_0 - r, x_0 + r) \subseteq I$. Therefore, we have

$$f(x) = f(x_0) + \frac{1}{2} f^{(2)}(x)(x-x_0)^2 \geq f(x_0)$$

for all $x \in (x_0 - r, x_0 + r)$ and thus, x_0 is a local minimum point of f as desired. □

Proposition 1.20. L'Hospital's Rule: *Let f and g be the differentiable functions on (a,b) and let $c \in (a,b)$ Assume that $f(c) = g(c) = 0$, in addition, $g'(x) \neq 0$ and $g(x) \neq 0$ for all $x \in (a,b) \setminus \{c\}$. If the limit $L := \lim_{x\to c} \frac{f'(x)}{g'(x)}$ exists, then so does $\lim_{x\to c} \frac{f(x)}{g(x)}$, moreover, we have $L = \lim_{x\to c} \frac{f(x)}{g(x)}$.*

Proof. Fix $c < x < b$. Then by the Cauchy Mean Value Theorem, there is a point $x_1 \in (c,x)$ such that

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(c)}{g(x) - g(c)} = \frac{f'(x_1)}{g'(x_1)}$$

$x_1 \in (c,x)$, so if $L := \lim_{x\to c} \frac{f'(x)}{g'(x)}$ exists, then $\lim_{x\to c+} \frac{f(x)}{g(x)}$ exists and is equal to L.

Similarly, we also have $\lim_{x\to c-} \frac{f(x)}{g(x)} = L$. The proof is finished. □

Proposition 1.21. *Let f be a function on (a,b) and let $c \in (a,b)$.*

(i) If $f'(c)$ exists, then the following limit exists (also called the symmetric derivatives of f at c):

$$f'(c) = \lim_{t\to 0} \frac{f(c+t) - f(c-t)}{2t}.$$

(ii) If $f^{(2)}(c)$ exists, then

$$f^{(2)}(c) = \lim_{t\to 0} \frac{f(c+t) - 2f(c) + f(c-t)}{t^2}.$$

Proof. For showing (i), note that we have

$$f'(c) = \lim_{t\to 0+} \frac{f(c+t) - f(c)}{t} = \lim_{t\to 0-} \frac{f(c+t) - f(c)}{t}.$$

Putting $t = -s$ into the second equality above, we see that

$$f'(c) = \lim_{s\to 0+} \frac{f(c-s) - f(c)}{-s}.$$

To sum up the two equations above, we have

$$f'(c) = \lim_{t\to 0+} \frac{f(c+t) - f(c-t)}{2t}.$$

Similarly, we have $f'(c) = \lim_{t\to 0-} \frac{f(c+t) - f(c-t)}{2t}$. Part (i) follows.

For showing Part (ii), let $h(t) := f(c+t) - 2f(c) + f(c-t)$ for $t \in \mathbb{R}$. Then $h(0) = 0$ and $h'(t) = f'(c+t) - f'(c-t)$. By using the L'Hospital's Rule and Part (i), we have

$$\lim_{t\to 0} \frac{f(c+t) - 2f(c) + f(c-t)}{t^2} = \lim_{t\to 0} \frac{h'(t)}{(t^2)'} = \lim_{t\to 0} \frac{f'(c+t) - f'(c-t)}{2t} = f^{(2)}(c).$$

The proof is complete. □

Definition 1.22. *A function f defined on (a,b) is said to be convex if for any pair $a < x_1 < x_2 < b$, we have*

$$f((1-t)x_1 + tx_2) \leq (1-t)f(x_1) + tf(x_2)$$

for all $t \in [0,1]$.

Proposition 1.23. *Let f be a C^2 function on (a,b). Then f is a convex function if and only if $f^{(2)}(x) \geq 0$ for all $x \in (a,b)$.*

Proof. For showing ($\Rightarrow$): assume that f is a convex function. Fix a point $c \in (a,b)$. f is convex, so we have $f(c) = f(\frac{1}{2}(c+t) + \frac{1}{2}(c-t)) \leq \frac{1}{2}f(c+t) + \frac{1}{2}f(c-t)$ for all $t \in \mathbb{R}$ with $c \pm t \in (a,b)$. By Proposition 1.21, we have

$$f^{(2)}(c) = \lim_{t \to 0} \frac{f(c+t) - 2f(c) + f(c-t)}{t^2}.$$

Therefore, we have $f^{(2)}(c) \geq 0$.
For ($\Leftarrow$), assume that $f^{(2)}(x) \geq 0$ for all $x \in (a,b)$. Fix $a < x_1 < x_2 < b$ and $t \in [0,1]$. Let $c := (1-t)x_1 + tx_2$. Then by the Lagrange Reminder Theorem, there are points $z_1 \in (x_1, c)$ and $z_2 \in (c, x_2)$ such that

$$f(x_2) = f(c) + f'(c)(x_2 - c) + \frac{1}{2}f^{(2)}(z_2)(x_2 - c)^2$$

and

$$f(x_1) = f(c) + f'(c)(x_1 - c) + \frac{1}{2}f^{(2)}(z_1)(x_1 - c)^2.$$

These two equations implies that

$$(1-t)f(x_1) + tf(x_2) = f(c) + (1-t)\frac{1}{2}f^{(2)}(z_1)(x_1 - c)^2 + t\frac{1}{2}f^{(2)}(z_2)(x_2 - c)^2 \geq f(c).$$

since $f^{(2)}(z_1)$ and $f^{(2)}(z_2)$ both are non-negative. Thus, f is convex. □

Corollary 1.24. *Let $p > 0$. The function $f(x) := x^p$ is convex on $(0, \infty)$ if and only if $p \geq 1$.*

Proof. Note that $f^{(2)}(x) = p(p-1)x^{p-2}$ for all $x > 0$. Then the result follows immediately from Proposition 1.23. □

Proposition 1.25. Netwon's Method*: Let f be a continuous real-valued function defined on $[a,b]$ with $f(a) < 0 < f(b)$ and $f(z) = 0$ for some $z \in (a,b)$. Assume that f is a C^2 function on (a,b) and $f'(x) \neq 0$ for all $x \in (a,b)$. Then there is $\delta > 0$ with $J := [z-\delta, z+\delta] \subseteq [a,b]$ which have the following property:*
if we fix any $x_1 \in J$ and let

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)} \tag{1.1}$$

for $n = 1, 2, ...$, then we have $z = \lim x_n$.

Proof. We first choose $r > 0$ such that $[z-r, z+r] \subseteq (a,b)$. We fix any point $x_1 \in (z-r, z+r)$ with $x_1 \neq z$. Then by the Lagrange Remainder Theorem, there is a point ξ between z and x_1 such that

$$0 = f(z) = f(x_1) + f'(x_1)(z - x_1) + \frac{1}{2}f^{(2)}(\xi)(z - x_1)^2.$$

This, together with Eq 1.1 above, we have

$$x_2 - x_1 = -\frac{f(x_1)}{f'(x_1)} = z - x_1 + \frac{f^{(2)}(\xi)}{2f'(x_1)}(z - x_1)^2.$$

Therefore, we have

$$x_2 - z = \frac{f^{(2)}(\xi)}{2f'(x_1)}(z - x_1)^2. \tag{1.2}$$

Note that the functions $f'(x)$ and $f^{(2)}(x)$ are continuous on $[z-r,z+r]$ and $f'(x)\neq 0$, hence, there is $M>0$ such that $|\frac{f^{2)}(u)}{2f'(v)}|\leq M$ for all $u,v\in[z-r,z+r]$. Then the Eq 1.2 implies that

$$|x_2-z|=|\frac{f^{(2)}(\xi)}{2f'(x_1)}(z-x_1)^2|\leq M(z-x_1)^2. \tag{1.3}$$

Choose $\delta>0$ such that $M\delta<1$ and $J:=[z-\delta,z+\delta]\subseteq(z-r,z+r)$. Note that Now we take any $x_1\in J$. Eq 1.3 implies that $|x_2-z|\leq M\cdot|z-x_1|^2\leq(M\delta)\cdot|x_1-z|$. By using Eq 1.1 inductively, we have a sequence (x_n) in J such that

$$|x_{n+1}-z|\leq M\cdot|z-x_n|^2\leq(M\delta)\cdot|x_n-z|$$

for all $n=1,2....$ Therefore, we have

$$|x_{n+1}-z|\leq(M\delta)^n\cdot|x_1-z|$$

for all $n=1,2...$, thus, $\lim x_n=z$. The proof is complete. □

Appendix: Differentiability on $\mathbb{R}^n$

Recall that for each element $x=(x_1,...,x_n)$ in $\mathbb{R}^n$, write $\|x\|:=\sqrt{|x_1|^2+\cdots+|x_n|^2}$ (call the *norm* of x). And for $a\in\mathbb{R}^n$ and $r>0$, put $B(a,r):=\{x\in\mathbb{R}^n:\|x-a\|<r\}$.

Lemma 1.26. *Every linear map on $\mathbb{R}^n$ is continuous.*

Proof. Let $T:\mathbb{R}^n\to\mathbb{R}^m$ be a linear map and let $\{e_1,...,e_n\}$ be the natural basis for $\mathbb{R}^n$. It suffices to show that the map T is continuous at 0 (why?). Let (x_i) be a sequence in $\mathbb{R}^n$ that converges to 0. If we write $x_i:=\sum_{k=1}^n t_i(k)e_k$, then $\lim\limits_{i\to\infty}t_i(k)=0$ for all $k=1,...,n$. This implies that $\lim\limits_{i\to\infty}T(x_i)=\sum\limits_{k=1}^n\lim\limits_{i\to\infty}t_i(k)Te_k=0$ as desired. □

Remark 1.27. *Notice that a linear map on an infinite dimensional space may not be continuous. For example, we consider an infinite dimensional vector space $E:=\bigcup_{n=1}^\infty\mathbb{R}^n$ whose norm is given by $\|x\|=\sum_{k=1}^\infty x(k)^2$ for $x=(x(k))_{k=1}^\infty\in E$. Define $T:E\to E$ by $Tx(k):=kx(k)$ for $k=1,2,....$ for $x\in E$. Then T is a linear map but it is discontinuous at 0 (why?).*
If you want to know more details about the infinite dimensional case, take the course of Functional Analysis in future.

Definition 1.28. *Let U be an open subset of $\mathbb{R}^n$ and let $f:U\to\mathbb{R}^m$ be a mapping. We say that f is differentiable at a point $a\in U$ if there is a (continuous) linear map $L(a):\mathbb{R}^n\to\mathbb{R}^m$ such that*

$$\lim_{v\to 0}\frac{\|f(a+v)-f(a)-L(a)(v)\|_{\mathbb{R}^m}}{\|v\|_{\mathbb{R}^n}}=0. \tag{1.4}$$

$L(a)$ is called a differential of f at a. f is said to be differentiable on U if it is differentiable at every point in U.

Proposition 1.29. *We keep the notation as given in Definition 1.28. Then we have the followings.*

(i) *f is differentiable at $a\in U$ if and only if there are a linear map $L(a):\mathbb{R}^n\to\mathbb{R}^m$ and a function $\alpha(a,\cdot):U\to\mathbb{R}^m$ such that*

$$f(x)=f(a)+L(a)(x-a)+\alpha(a,x)\quad \textit{for all } x\in U \textit{ and }\quad \lim_{x\to a}\frac{\|\alpha(a,x)\|}{\|x-a\|}=0. \tag{1.5}$$

(ii) If f is differentiable at a, then f is continuous at a.
(iii) A differential of f at $a \in U$ is unique if it exists.

From now on, we write $f'(a)$ for the differential of f at a.

Proof. For Part $(i)(\Rightarrow)$, if f is differentiable at a, then we put

$$\alpha(a,x) := f(x) - f(a) - L(a)(x-a)$$

for $x \in U$. Then Eq 1.4 implies that $\lim_{x\to a} \frac{\|\alpha(a,x)\|}{\|x-a\|} = 0$ as desired. The converse is clear.
For Part (ii), we keep the notation as in Part (i). Since $\lim_{x\to a} \frac{\|\alpha(a,x)\|}{\|x-a\|} = 0$, we have $\lim_{x\to a} \|\alpha(a,x)\| = 0$. Thus, $\lim_{x\to a}(f(x) - f(a)) = 0$ by Eq 1.5 because every linear map is continuous. For showing (iii), let $L_1(a)$ and $L_2(a)$ be the linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$. Let $\alpha_1(a,\cdot)$ and $\alpha_2(a,\cdot)$ be the functions given as in Part (i). From this we have

$$L_1(a)(x-a) + \alpha_1(a,x) = L_2(a)(x-a) + \alpha_2(a,x)$$

for all $x \in U$. Now choose $r > 0$ such that $B(a,r) \subseteq U$ and so we have $L_1(a)(v) + \alpha_1(a, a+v) = L_2(a)(v) + \alpha_2(a, a+v)$ for all $v \in B(0,r)$. Now if we fix $0 \neq v \in B(0,r)$, then we have

$$L_1(a)(tv) + \alpha_1(a, a+tv) = L_2(a)(tv) + \alpha_2(a, a+tv)$$

for all $0 < t \leq 1$. From this, taking $t \to 0+$, we have $L_1(a)(\frac{v}{\|v\|}) = L_2(a)(\frac{v}{\|v\|})$ and thus, $L_1(a)(v) = L_2(a)(v)$ for all $0 \neq v \in B(0,r)$. Then by the linearity of $L_1(a)$ and $L_2(a)$, we conclude that $L_1(a)(v) = L_2(a)(v)$ for all $v \in \mathbb{R}^n$. The proof is complete. □

Proposition 1.30. Chain Rule: *Let $f : U \to V$ and $g : V \to \mathbb{R}^l$ be the mappings where U and V are the open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. Let $a \in U$ and put $b := f(a)$. If $f'(a)$ and $g'(b)$ both exist, then $(g \circ f)'(a)$ exists and $(g \circ f)'(a) = g'(b) \circ f'(a) : \mathbb{R}^n \to \mathbb{R}^l$.*

Proof. Put $y = f(x)$. Let $\alpha(a,\cdot) : U \to \mathbb{R}^n$ and $\beta(b,\cdot) : V \to \mathbb{R}^l$ be the functions given as in Proposition 1.29 above. Notice that we have

$$f(x) = f(a) + f'(a)(x-a) + \alpha(a,x)$$

for all $x \in U$ and

$$g(y) = g(b) + g'(b)(y-b) + \beta(b,y)$$

for all $y \in V$. From this we have

$$\begin{aligned} g \circ f(x) &= g \circ f(a) + g'(b)(f(x) - f(a)) + \beta(f(a), f(x)) \\ &= g \circ f(a) + g'(b)f'(a)(x-a) + g'(b)(\alpha(a,x)) + \beta(f(a), f(x)) \end{aligned}$$

for all $x \in U$. Let

$$\gamma(a,x) := g'(b)(\alpha(a,x)) + \beta(f(a), f(x))$$

for $x \in U$. Then by Proposition 1.29, we need to show that

$$\lim_{x\to a} \frac{\|\gamma(a,x)\|}{\|x-a\|} = 0.$$

Since $\lim_{x\to a} \frac{\alpha(a,x)}{\|x-a\|} = 0$ and every linear map is continuous , we have $\lim_{x\to a} g'(b)(\frac{\alpha(a,x)}{\|x-a\|}) = 0$. Hence, it suffices to show that $\lim_{x\to a} \frac{\beta(b,y)}{\|x-a\|} = 0$.
In fact, let $\varepsilon > 0$, then by the construction of $\beta(b,y)$, there is $\delta_1 > 0$ such that

$$\frac{\|\beta(b,y)\|}{\|b-y\|} < \varepsilon \quad \text{whenever} \quad 0 < \|y-b\| < \delta_1.$$

Since f is continuous at a, there is $\delta_2 > 0$ such that $\|y - b\| < \delta_1$ whenever $0 < \|x - a\| < \delta_2$. On the other hand, we have

$$\frac{b-y}{\|x-a\|} = f'(a)(\frac{x-a}{\|x-a\|}) + \frac{\alpha(a,x)}{\|x-a\|}.$$

for all $x \in U\backslash\{a\}$. Since $f'(a) : \mathbb{R}^n \to \mathbb{R}^m$ is continuous and the unit sphere $S_{n-1} := \{v \in \mathbb{R}^n : \|v\| = 1\}$ is compact, we have

$$\|f'(a)(\frac{x-a}{\|x-a\|})\| \le \sup_{v\in S_{n-1}} \|f'(a)(v)\| < \infty$$

for all $x \in U\backslash\{a\}$. Also, there is $0 < \delta < \delta_2$ such that $x \in U$ and $\frac{\|\alpha(a,x)\|}{\|x-a\|} < 1$ as $0 < \|x-a\| < \delta$. Thus, there is $M > 0$ such that $\frac{\|b-y\|}{\|x-a\|} \le M$ whenever $0 < \|x - a\| < \delta$. This implies that if $y = f(x) \neq b$ and $0 < \|x - a\| < \delta$, then we have

$$\frac{\|\beta(b,y)\|}{\|x-a\|} = \frac{\|\beta(b,y)\|}{\|b-y\|}\frac{\|b-y\|}{\|x-a\|} \le \varepsilon M.$$

Notice that $\beta(b, y) = 0$ if $y = b$. Therefore, if $0 < \|x - a\| < \delta$, then we have

$$\frac{\|\beta(b,y)\|}{\|x-a\|} \le \varepsilon M.$$

The proof is complete. □

To end this appendix, we are going to define the higher order differentials of f. Before giving the definition, let us recall the notation of multilinear maps. Let E and F be vector spaces. A mapping $T : E \times \cdots \times E(r\text{-copies}) \to F$ is called a r-linear map if T is linear for each variable, more precisely, for $1 \le k \le r$ and $x_1, ..., x_{k-1}, x_{k+1}, ..., x_r \in E$, the map $x \in E \mapsto T(x_1, ..., x_{k-1}, x, x_{k+1}, ..., x_r) \in F$ is linear. Write $L^{(r)}(E, F)$ for the set of all r-linear maps. Clearly, $L^{(r)}(E, F)$ is a vector space.

Lemma 1.31. *$L^{(r)}(\mathbb{R}^n, \mathbb{R}^m) = \mathbb{R}^{n^r m}$ for $r = 1, 2, ...$ Consequently, the space $L^{(r)}(\mathbb{R}^n, \mathbb{R}^m)$ have the norm structure induced by $\mathbb{R}^{n^r m}$.*

Proof. Clearly, we have $L^{(1)}(\mathbb{R}^n, \mathbb{R}^m) = M_{m\times n}(\mathbb{R}) = \mathbb{R}^{nm}$. Notice that we have $L^{(2)}(\mathbb{R}^n, \mathbb{R}^m) = L^{(1)}(\mathbb{R}^n, L^{(1)}(\mathbb{R}^n, \mathbb{R}^m))$ and so, $L^{(2)}(\mathbb{R}^n, \mathbb{R}^m) = \mathbb{R}^{n^2 m}$. Using induction on r, we see that $L^{(r)}(\mathbb{R}^n, \mathbb{R}^m) = \mathbb{R}^{n^r m}$. □

Definition 1.32. *We keep the notation as in Definition 1.28. Notice that if f is differentiable on U, then the differential of f gives a map*

$$f' : a \in U \mapsto f'(a) \in L^{(1)}(\mathbb{R}^n, \mathbb{R}^m).$$

Note that the space $L^{(1)}(\mathbb{R}^n, \mathbb{R}^m)$ have the natural norm structure given by Lemma 1.31, that is, $L^{(1)}(\mathbb{R}^n, \mathbb{R}^m) = \mathbb{R}^{nm}$. If f' is differentiable on U in the sense of Definition 1.28, then for each $a \in U$, it is naturally led to define

$$f^{(2)}(a) := (f')'(a) \in L^{(1)}(\mathbb{R}^n, L^{(1)}(\mathbb{R}^n, \mathbb{R}^m)) = L^{(2)}(\mathbb{R}^n, \mathbb{R}^m) = \mathbb{R}^{n^2 m}.$$

Thus, one can define inductively the r-th differential of f at a as the following

$$f^{(r)}(a) := (f^{r-1})'(a) \in L^{(r)}(\mathbb{R}^n, \mathbb{R}^m).$$

2. 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 2.1. *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 2.2. *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.

$$L(f,Q) - L(f,P) = m_s(f,Q)(c - x_{s-1}) + m_{s+1}(f,Q)(x_s - c) - m_s(f,P)(x_s - x_{s-1}) \geq 0. \tag{2.1}$$

Now by considering $-f$ in the Inequality 2.1 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. □

The following notion plays an important role in this chapter.

Definition 2.3. *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.

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

Notice that the upper integral and lower integral of f must exist by Remark 2.1.

Remark 2.4. Appendix: We call a partially set $(I, \leq)$ a *directed set* if for each pair of elements i_1 and i_2 in I, there is $i_3 \in I$ such that $i_1 \leq i_3$ and $i_2 \leq i_3$.
A *net* in $\mathbb{R}$ is a real-valued function f defined on a directed set I, write $f = (x_i)_{i\in I}$, where $x_i := f(i)$ for $i \in I$.
We say that a net (x_i) converges to a point $L \in \mathbb{R}$ (call a limit of (x_i)) if for any $\varepsilon > 0$, there is $i_0 \in I$ such that $|x_i - L| < \varepsilon$ for all $i \geq i_0$.
Using the similar argument as in the sequence case, a limit of (x_i) is unique if it exists and we write $\lim_i x_i$ for its limits.

Example 2.5. Appendix: Using the notation given as before, let

$$I := \{P : \ P \text{ is a partitation on } [a,b]\ \}.$$

We say that $P_1 \leq P_2$ for $P_1, P_2 \in I$ if $P_1 \subseteq P_2$. Clearly, I is a directed set with this order. If we put $u_P := U((f,P)$, then we have

$$\lim_P u_P = \overline{\int_a^b} f.$$

In fact, let $\varepsilon > 0$. Then by the definition of an upper integral, there is $P_0 \in I$ such that

$$\overline{\int_a^b} f \leq U(f,P_0) \leq \overline{\int_a^b} f + \varepsilon.$$

Lemma 2.2 tells us that whenever $P \in I$ with $P \geq P_0$, we have $U(f,P) \leq U(f,P_0)$. Thus we have $|u_P - \overline{\int_a^b} f| < \varepsilon$ whenever $P \geq P_0$ as desired.

Proposition 2.6. *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 2.2 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 2.2, 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

$$\underline{\int_a^b} f + \underline{\int_a^b} g \leq \underline{\int_a^b} (f+g). \tag{2.2}$$

As before, we consider $-f$ and $-g$ in the Inequality 2.2, 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 2.6 (*iii*) may hold in general.

Example 2.7. *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 2.8. *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 2.9. *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 2.6 (*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 2.10. *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*

$$0 \leq U(f,P) - L(f,P) = \sum_{i=1}^{n} \omega_i(f,P)\Delta x_i < \varepsilon. \tag{2.3}$$

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 2.3 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 2.11. *Theorem 2.10 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. See the Appendix 3 below for details.*

Example 2.12. *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 2.10, 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. $\square$

Proposition 2.13. *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 2.14. *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 2.10 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. $\square$

Lemma 2.15. *Let g be a convex function defined on $[a,b]$. Then for $a<c<x<d<b$, we have*

$$\frac{g(x)-g(c)}{x-c} \leq \frac{g(d)-g(x)}{d-x}.$$

Consequently, if $a<x_1<x_2<x_3<x_4<b$, we have

$$\frac{g(x_2)-g(x_1)}{x_2-x_1} \leq \frac{g(x_4)-g(x_3)}{x_4-x_3}. \tag{2.4}$$

Proof. Let $\ell(x)$ be the straight line between the points $(c,g(c))$ and $(d,g(d))$. Then we have $g(x)\leq \ell(x)$ for all $x\in[c,d]$ by the convexity. This implies the following that we desired.

$$\frac{g(x)-g(c)}{x-c} \leq \frac{\ell(x)-\ell(c)}{x-c} = \frac{\ell(d)-\ell(x)}{d-x} \leq \frac{g(d)-g(x)}{d-x}.$$

$\square$

Proposition 2.16. (Jensen's inequality): *Let $g:[a',b']\longrightarrow \mathbb{R}$ be a convex function and $f\in R([0,1])$ such that $f([0,1])\subseteq[a,b]\subseteq(a',b')$ and $g\circ f\in R([0,1])$. Then we have*

$$g(\int_0^1 f(x)dx) \leq \int_0^1 (g\circ f)(x)dx.$$

Proof. Notice that if we let $c := \int_0^1 f$, then $c\in[a,b]$ and hence, $g(c)$ is defined. Notice that by Eq2.4 above the set $\{\frac{g(c)-g(x)}{c-x} : a'<x<c\}$ is bounded above and so, $s:=\sup\{\frac{g(c)-g(y)}{c-y} : a'<y<c\}$ is defined. Thus, we have

$$g(c)-g(y)\leq s(c-y) \quad \text{for all } a'<y<c.$$

On the other hand, if $c<y_1<b'$, then by Eq2.4 again we have

$$\frac{g(c)-g(y)}{c-y} \leq \frac{g(y_1)-g(c)}{y_1-c} \quad \text{for all } a'<y<c.$$

Hence, we have $s\leq \frac{g(y_1)-g(c)}{y_1-c}$ for all $c<y_1<b'$. Thus, we have

$$g(c)-g(y)\leq s(c-y) \quad \text{for all } a'<y<b' \text{ with } y\neq c. \tag{2.5}$$

Note that Eq 2.5 clearly holds for $y=c$. Thus, Eq2.5 is true for all $a'<y<b'$. Now if put $y=f(x)$, then we have $g(c)+s(f(x)-c)\leq (g\circ f)(x)$ for all $x\in[0,1]$. This gives

$$g(c)=g(c)+s\int_0^1 (f(x)-c)dx \leq \int_0^1 (g\circ f)(x)dx.$$

The proof is complete. $\square$

Example 2.17. *Let $a_1,...,a_n$ be any real numbers. Let $p>1$. Then we have*

$$(\frac{|a_1|+\cdots|a_n|}{n})^p \leq \frac{1}{n}\sum_{k=1}^n |a_k|^p.$$

To see this, , the results obtained by applying the Jensen's inequality for the convex function $g(x)=x^p$ for $x\geq 0$ and $f(t):=|a_k|$ for $t\in[(k-1)/n,k/n)$ for $k=1,...,n$.

Proposition 2.18. *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*

$$\int_a^b f = \int_a^c f + \int_c^b f. \tag{2.6}$$

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 2.10. 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 2.6 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 2.10. □

Proposition 2.19. *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 2.10, 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. □

Remark 2.20. *In the proof of Proposition 2.19, 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 2.21. *Assume that $f : [a,b] \longrightarrow [c,d]$ is integrable and $g : [c,d] \longrightarrow \mathbb{R}$ is continuous. Then the composition $g \circ f \in R[a,b]$.*

Proof. Let $\varepsilon > 0$. Note that g is uniformly continuous on $[c,d]$ because g is continuous on $[c,d]$. Then there is $\delta > 0$ such that $|g(y) - g(y')| < \varepsilon$ whenever $y, y' \in [c,d]$ with $|y - y'| < \delta$. On the other hand, since $f \in R[a,b]$, there is a partition P on $[a,b]$ such that $\sum \omega_k(f,P)\Delta x_k < \varepsilon\delta$. Hence, we have

$$\delta \sum_{k:\omega_k(f,P)\ge\delta} \Delta x_k \le \delta \sum_{k:\omega_k(f,P)\ge\delta} \omega_k(f,P)\Delta x_k < \varepsilon\delta.$$

This implies that

$$\sum_{k:\omega_k(f,P)\ge\delta} \Delta x_k < \varepsilon.$$

On the other hand, by the choice of δ, we see that $\omega_k(g\circ f,P)<\varepsilon$ whenever $\omega_k(f,P)<\delta$. Therefore, we can conclude that

$$\sum_k \omega_k(g\circ f,P)\Delta x_k=\sum_{k:\omega_k(f,P)<\delta}\omega_k(g\circ f,P)\Delta x_k+\sum_{k:\omega_k(f,P)\geq\delta}\omega_k(g\circ f,P)\Delta x_k<\epsilon(b-a)+2M\epsilon$$

where $M:=\sup|f(x)|$. The proof is complete. □

Remark 2.22. *The composition of integrable functions need not be integrable. For example, if we put f is given as in Example 2.12 and $g(x)=x$ for $x=1/n, n=1,2,...$; otherwise $g(x)=0$. Then $f,g\in R[0,1]$ but $g\circ f\notin R[0,1]$.*

Proposition 2.23. (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

$$\int_a^b f(x)g(x)dx=f(\xi)\int_a^b g(x)dx. \tag{2.7}$$

In particular, there is a point ξ in (a,b) such that $f(\xi)=\frac{1}{b-a}\int_a^b f(x)dx$.

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 $a\leq x_1<x_2\leq b$. From this, since $g\leq 0$, we have

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

for all $x\in[a,b]$. From this and Proposition 2.19 above, we have

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

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

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

Therefore, there is a point $\xi\in[x_1,x_2]\subseteq[a,b]$ so that the Equation 2.7 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 $a\leq x_1<x_2\leq b$ 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

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

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. □

Example 2.24. *We have* $\lim_n \int_0^{\pi/2} \sin^n x dx = 0$. *To see this, for any* $0 < \varepsilon < \pi/2$ *and for each* $n = 1, 2...$, *the Mean value theorem gives a point* $\xi_n \in (0, \frac{\pi}{2} - \varepsilon)$ *such that*

$$\begin{aligned} 0 < \int_0^{\pi/2} \sin^n x dx &= (\int_0^{\frac{\pi}{2}-\varepsilon} + \int_{\frac{\pi}{2}-\varepsilon}^{\pi/2}) \sin^n x dx \\ &\leq \sin^{n-1}\xi_n \int_0^{\frac{\pi}{2}-\varepsilon} \sin x dx + \int_{\frac{\pi}{2}-\varepsilon}^{\pi/2} \sin^n x dx \\ &< \sin^{n-1}(\frac{\pi}{2} - \varepsilon) + \varepsilon. \end{aligned}$$

Taking $n \to \infty$, *we have* $\overline{\lim}_n \int_0^{\pi/2} \sin^n x dx = 0$. *The proof is finished.*

Now if $f \in R[a,b]$, then by Proposition 2.18, we can define a function $F : [a,b] \to \mathbb{R}$ by

$$F(c) = \begin{cases} 0 & \text{if } c = a \\ \int_a^c f & \text{if } a < c \leq b. \end{cases} \tag{2.8}$$

Theorem 2.25. 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* F *on* $[a,b]$ *which is differentiable on* (a,b) *with* $F' = f$, *then* $\int_a^b f = F(b) - F(a)$. *In this case,* F *is called an indefinite integral of* f. *(**note**: if* F_1 *and* F_2 *both are the indefinite integrals of* f, *then by the Mean Value Theorem, we have* $F_2 = F_1 +$ *constant).*
- *(ii) The function* F *defined as in Eq. 2.8 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 $\xi_i \in (x_{i-1}, x_i)$ such that $F(x_i) - F(x_{i-1}) = F'(\xi_i)\Delta x_i = f(\xi_i)\Delta x_i$. So, we have

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

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. □

Definition 2.26. *For each function f on $[a,b]$ and a partition $P : a = x_0 < \cdots < x_n = b$, we call $R(f,P,\{\xi_i\}) := \sum_{i=1}^N f(\xi_i)\Delta x_i$, 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 $A = \lim_{\|P\|\to 0} R(f,P,\{\xi_i\})$, 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 2.27. *Let f be a function defined on $[a,b]$. If the limit $\lim_{\|P\|\to 0} R(f,P,\{\xi_i\}) = A$ 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 2.28. *$f \in R[a,b]$ 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 2.10.
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 2.29. *$f \in R[a,b]$ 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 2.28 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 $\xi_i \in [x_{i-1}, x_i]$.
Note that f is automatically bounded in this case by Proposition 2.27.
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

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

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

$$\overline{\int_a^b} f(x)dx \le U(f,P) \le A + \varepsilon(1 + b - a). \tag{2.9}$$

By considering $-f$, note that the Riemann sum of $-f$ will converge to $-A$. The inequality 2.9 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 complete. □

Proposition 2.30. *Let $f \in C[c,d]$. Let $\phi : [a,b] \longrightarrow [c,d]$ be a function with $\phi(a) = c$ and $\phi(b) = d$. Assume that ϕ is a C^1 function over $[a,b]$, that is, ϕ' can be extended to a continuous function on $[a,b]$. Then we have*

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

Proof. Notice that since f is continuous on $[c,d]$, the Fundamental Theorem of Calculus yields an indefinite integral F of f on $[c,d]$. Put $h(t) := F \circ \phi(t)$ for $t \in [a,b]$. Then by the chain rule, we see that $h'(t) = F'(\phi(t)) \cdot \phi'(t) = f(\phi(t)) \cdot \phi'(t)$ for $t \in (a,b)$. Using the Fundamental Theorem of Calculus again, we have

$$\int_a^b f(\phi(t)) \cdot \phi'(t)dt = \int_a^b h'(t)dt = h(b) - h(a) = F(d) - F(c) = \int_c^d f(x)dx.$$

The proof is finished. □

The following theorem shows us that the assumption of the continuity of f in Proposition 2.30 can be replaced by a weaker condition.

Theorem 2.31. (Change of variable formula): *Let $f \in R[c,d]$. Let $\phi : [a,b] \longrightarrow [c,d]$ be a C^1 function over $[a,b]$ with $\phi(a) = c$ and $\phi(b) = d$ satisfying $\phi' > 0$. 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 using Theorem 2.29, 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 2.28 and Theorem 2.29, there is $\delta_1 > 0$ such that

$$|A - \sum f(\eta_k)\triangle x_k| < \varepsilon \tag{2.10}$$

and

$$\sum \omega_k(f,P)\triangle x_k < \varepsilon \tag{2.11}$$

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 $t \in [a,b]$.
Note that 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

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

Now for any $\xi_k \in [t_{k-1}, t_k]$, we have

$$\begin{aligned} |A - \sum f(\phi(\xi_k))\phi'(\xi_k)\triangle t_k| \le& |A - \sum f(\phi(\xi_k^*))\phi'(\xi_k^*)\triangle t_k| \\ &+ |\sum f(\phi(\xi_k^*))\phi'(\xi_k^*)\triangle t_k - \sum f(\phi(\xi_k))\phi'(\xi_k^*)\triangle t_k| \\ &+ |\sum f(\phi(\xi_k))\phi'(\xi_k^*)\triangle t_k - \sum f(\phi(\xi_k))\phi'(\xi_k)\triangle t_k| \end{aligned} \tag{2.12}$$

Notice that inequality 2.10 implies that

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

On the other hand, we have

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

Concerning about the last inequality in 2.12, 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 $|f(x)| \leq M$ for all $x \in [c, d]$.

Finally by inequality 2.12, we have

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

Finally, we have to show that $f \circ \phi \in R[a, b]$. To see this, we have shown that the function $f \circ \phi(t)\phi'(t) \in R[a, b]$ by above. Since $\phi' > 0$ is continuous on $[a, b]$, $\frac{1}{\phi'}$ is continuous on $[a, b]$ and thus $\frac{1}{\phi'} \in R[a, b]$. This implies that the function $f \circ \phi = \frac{1}{\phi'}(f \circ \phi \cdot \phi') \in R[a, b]$ as desired. The proof is complete. □

Definition 2.32. *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.

Clearly, the Cauchy criterion will imply the following immediately.

Proposition 2.33. *Let $f : [a, \infty) \longrightarrow \mathbb{R}$ be a function given as in Definition 2.32.*

(i) The improper integral $\int_a^\infty f$ exists if and only if for any $\varepsilon > 0$, there is $M > 0$ such that $|\int_A^B f| < \varepsilon$ whenever $M < A < B$.

(ii) Let g be a non-negative function defined on $[a, \infty)$ such that $|f| \leq g$ on $[a, \infty)$. If $\int_a^\infty g$ is convergent, then so is $\int_a^\infty f$.

(iii) Suppose that $0 \leq g \leq f$ on $[a, \infty)$. If $\int_a^\infty g$ is divergent, then so is $\int_a^\infty f$.

Similar assertion holds when f is defined on $(a, b]$.

Remark 2.34. *By using the Cauchy Theorem,it is clear that if $\int_a^\infty |f|$ is convergent, then so is the integral $\int_a^\infty f$. However, the converse does not hold. It is quit different from the case when f defined on $[a, b]$.*
For example, if $f(x) = \frac{(-1)^{n-1}}{n}$ as $n \in [n-1, n)$ $n = 1, 2, ...,$ then $\int_a^\infty f$ is convergent (it will be shown in the last chapter) but $\int_a^\infty |f|$ is divergent.

Example 2.35. *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 $I(s) = \lim_{\eta\to 0+} \int_\eta^1 x^{s-1}e^{-x}dx$ 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. □

3. Appendix: Lebesgue integrability theorem

Throughout this section, let f be a $\mathbb{R}$-valued function defined on $[a, b]$ and let $M := \sup|f(x)|$.

Definition 3.1. *A subset A of $\mathbb{R}$ is said to have measure zero (or null set) if for every $\varepsilon > 0$, there is a sequence of open intervals, (a_n, b_n) such that $A \subseteq \bigcup(a_n, b_n)$ and $\sum(b_n - a_n) < \varepsilon$.*

Clearly we have the following assertion.

Lemma 3.2. *If (A_n) is a sequence of null sets, then so is $\bigcup A_n$. Consequently, all countable sets are null sets.*

From now on, we use the following notation in the rest of this section.

(1) For each subset A of $\mathbb{R}$, put $\omega(f, A) := \sup\{|f(x) - f(x')| : x, x' \in A\}$.
(2) For $c \in [a, b]$, put $\omega(f, c) := \inf\{\omega(f, B(c, r)) : r > 0\}$, where $B(c, r) := (c - r, c + r)$.

The following is easy shown directly from the definition.

Lemma 3.3. *The function f is continuous at $c \in [a, b]$ if and only if $\omega(f, c) = 0$.*

Theorem 3.4. Lebesgue integrability theorem: *Retains the notation as above. Let $D := \{c \in [a, b] : f$ is discontinuous at $c\}$. Then $f \in R[a, b]$ if and only if D has measure zero.*

Proof. For each positive integer n, let $D_n := \{x \in [a, b] : \omega(f, x) \geq \frac{1}{n}\}$. Then we have $D = \bigcup_{n=1}^\infty D_n$.
For ($\Rightarrow$), assume that $f \in R[a, b]$. Then by Lemma 3.2, it suffices to show that each D_n is a null set.

Fix a positive integer m such that $D_m \neq \emptyset$. Now Let $\varepsilon > 0$. Since $f \in R[a,b]$, there is a partition $P : a = x_0 < \cdots < x_n = b$ such that $\sum \omega_k(f,P)\Delta x_k < \frac{\varepsilon}{m}$. Notice that $c \in D_m$ if and only if $\omega(f, B(c,\delta)) \geq \frac{1}{m}$ for all $\delta > 0$, where $B(c,\delta) := (c-\delta, c+\delta)$. Thus, if $(x_{k-1}, x_k) \cap D_m \neq \emptyset$, then $\omega_k(f,P) \geq \frac{1}{m}$. This implies that

$$\begin{aligned}\frac{\varepsilon}{m} &> \sum_{k=1}^{n} \omega_k(f,P)\Delta x_k \\ &\geq \sum_{k:(x_{k-1},x_k)\cap D_m \neq \emptyset} \omega_k(f,P)\Delta x_k \\ &\geq \frac{1}{m} \sum_{k:(x_{k-1},x_k)\cap D_m \neq \emptyset} \Delta x_k.\end{aligned}$$

Therefore, we have $D_m \subseteq \bigcup_{k:(x_{k-1},x_k)\cap D_m \neq \emptyset} [x_{k-1}, x_k]$ and

$$\sum_{k:(x_{k-1},x_k)\cap D_m \neq \emptyset} \Delta x_k < \varepsilon.$$

Thus, D_m is a null set for each positive integer m as desired.
Now for showing $(\Leftarrow)$, assume that the set D of all discontinuous points of f is a null set.
We first claim that each D_m is a closed set. To see this, note that a point $c \in D_m$ if and only if $\omega(f, B(c,r)) \geq \frac{1}{m}$ for all $r > 0$ if and only if for all $\eta > 0$ and for all $r > 0$, there are points $x', x'' \in B(c,r)$ such that $|f(x') - f(x'')| > \frac{1}{m} - \eta$. Now let (c_n) be a sequence in D_m converging to a point c. Let $r > 0$ and $\eta > 0$. Then there is c_N such that $|c_N - c| < \frac{r}{2}$. Since $c_N \in D_m$, there are $x', x'' \in B(c_N, \frac{r}{2})$ such that $|f(x') - f(x'')| > \frac{1}{m} - \eta$. Since $x', x'' \in B(c_N, \frac{r}{2})$, $x', x'' \in B(c,r)$. Thus, $c \in D_m$ is as desired. This shows that D_m is a closed subset of $[a,b]$, and hence it is compact.
Let $\varepsilon > 0$ and let m be a positive integer such that $1/m < \varepsilon$. By the assumption $D = \bigcup_{l=1}^{\infty} D_l$ is a null set and so is the set D_m. Then there is a sequence of open intervals, say $\{(a_j, b_j)\}$, such that $D_m \subseteq \bigcup (a_j, b_j)$ and $\sum (b_j - a_j) < \varepsilon$. Since D_m is compact, there are finitely many (a_j, b_j)'s for $j = 1, ..., K$ such that $D_m \subseteq \bigcup_{j=1}^{K} (a_j, b_j)$. Note that we may assume that the sequence $a_1 < b_1 < a_2 < b_2 < \cdots < a_K < b_K$. Choose a partition $Q := (\{a_j, b_j : j = 1, ..., K\} \cup \{a, b\}) \cap [a,b]$ on $[a,b]$ and rewrite Q as $a = x_0 < \cdots < x_n = b$.
Put $I := \{j : [x_{j-1}, x_j] \cap D_m = \emptyset\}$ and $II := \{j : [x_{j-1}, x_j] \cap D_m \neq \emptyset\}$.
Note that if $j \in I$, then $\omega(f,x) < \frac{1}{m}$ for all $x \in [x_{j-1}, x_j]$. Hence, for each $x \in [x_{j-1}, x_j]$, there is $\delta_x > 0$ such that $\omega(f, B(x, \delta_x)) < \frac{1}{m}$. Then by the compactness of $[x_{j-1}, x_j]$, there is a partition $P'_j : x_{j-1} = x'_0 < \cdots < x'_l = x_j$ on $[x_{j-1}, x_j]$ such that $\omega_{j'}(f, P'_j) < \frac{1}{m}$ for all $j' = 1, ..., l$. Thus, we have $\sum_{j'} \omega_{j'}(f, P'_j)\Delta x_{j'} < \frac{1}{m}(x_j - x_{j-1}) < \varepsilon(x_{j-1} - x_j)$ whenever $j \in I$.
On the other hand, if $j \in II$, then $[x_{j-1}, x_j] \cap D_m \neq \emptyset$. Since $\sum_{j=1}^{K} (b_j - a_j) < \varepsilon$, we see that $\sum_{j \in II} \omega j(f,Q)\Delta x_j < 2M\varepsilon$.
Now put $P := Q \cup \bigcup_{j \in I} P'_j : a = y_0 < \cdots < y_N = b$. From the above argument, we have shown that $\sum_{i=1}^{N} \omega_i(f,P)\Delta y_i < \varepsilon(b-a) + 2M\varepsilon$. Thus $f \in R[a,b]$. The proof is complete. □

4. Some results of sequences of functions

Proposition 4.1. *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

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

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

$$f(x) = \int_c^x g(t)dt + f(c). \tag{4.1}$$

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 4.1 will tell us that f' exists and $f' = g$ on (a,b). The proof is finished. □

Proposition 4.2. *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

$$f_m(x) - f_n(x) = f_m(c) - f_n(c) + (f_m'(\xi) - f_n'(\xi))(x - c). \tag{4.2}$$

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

$$|f_m'(x) - f_n'(x)| < \varepsilon \tag{4.3}$$

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.4.3 implies that

$$|\frac{f_m(x) - f_m(u)}{x-u} - \frac{f_N(x) - f_N(u)}{x-u}| \leq \varepsilon \tag{4.4}$$

for all $m \geq N$ and for all $x \in (a,b)$ with $x \neq u$.
Taking $m \to \infty$ in Eq.4.4, 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

$$|\frac{f(x) - f(u)}{x-u} - f_N'(u)| \leq 2\varepsilon \tag{4.5}$$

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.4.5 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 4.3. *The uniform convergence assumption of (f_n') in the Propositions above is essential.*

Example 4.4. *Let $f_n(x) := \frac{x}{1+n^2x^2}$ for $x \in (-1,1)$. Then we have*

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

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 4.1 and 4.2 does not hold. Note that (f_n') does not converge uniformly to g on $(-1,1)$.

Proposition 4.5. *(***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 $f_n(x) \leq f_{n+1}(x)$ 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

$$g_n(x_n) \geq \varepsilon. \tag{4.6}$$

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

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

because $g_m(x_{n_i}) \downarrow 0$ as $m \to \infty$. This contradicts to the Inequality 4.6.
Method II: Let $\varepsilon > 0$. Fix $x \in A$. Since $g_n(x) \downarrow 0$, there is $N(x) \in \mathbb{N}$ such that $0 \leq g_n(x) < \varepsilon$ for all $n \geq N(x)$. 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 $1 \leq i \leq m$. This implies that

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

for all $n \geq N \geq N(x_i)$. □

5. Absolutely convergent series

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

Definition 5.1. *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 5.2. 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 5.3. *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 5.4. *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 5.5. *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 $M > \max\{j : 1 \leq \sigma(j) \leq N\}$. Then $\sigma(i) \geq N$ if $i \geq M$. This together with $(*)$ imply that if $i \geq M$ 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. □

In view of Theorem 5.5, it is naturally to introduce the following definition.

Definition 5.6. *A series $\sum x_n$ is said to be unconditionally convergent if whenever π is a bijection on $\mathbb{Z}_+$ the series $\sum_n x_{\pi(n)}$ is convergent.*

Theorem 5.7. *Let $\sum_n x_n$ be a series of numbers. Then the following are equivalent.*

(i) $\sum_n x_n$ is unconditionally convergent.
(ii) For any subsequence of positive integers $n_1 < n_2 < \cdots$, the series $\sum_k x_{n_k}$ is convergent.
(iii) For any choice of sign sequence (ε_n), that is $\varepsilon_n = \pm 1$, the series $\sum_n \varepsilon_n x_n$ is convergent.
(iv) For any $\varepsilon > 0$, there is a positive integer N such that $|\sum_{i \in A} x_i| < \varepsilon$ whenever A is a finite subset of $\mathbb{Z}_+$ with $N < \min A$.

Proof. The route of the proof is as the following.

$$(i) \Rightarrow (iv) \Rightarrow (ii) \Rightarrow (iv) \Rightarrow (i); \quad \text{and } (ii) \Leftrightarrow (iii).$$

Part $(ii) \Leftrightarrow (iii)$ is clear.
For showing $(i) \Rightarrow (iv)$. Assume that (iv) does not hold. Hence, there is $\varepsilon > 0$ and there is a sequence of finite subsets (A_n) of $\mathbb{Z}_+$ such that $\max A_n < \min A_{n+1}$ and $|\sum_{i \in A_n} x_i| \ge \varepsilon$ for all n. From this we see that $A_n \cap A_m = \emptyset$ for all $m \ne n$ and there is a bijection $\pi : \mathbb{Z}_+ \to \mathbb{Z}_+$ such that each $A_n = \{\pi(i_n) < \pi(i_n + 1) < \cdots < \pi(i_n + p_n)\}$ for some positive integers i_n and p_n. Then by the construction of A_n, the series $\sum_n x_{\pi(n)}$ is divergent, hence (i) does not hold.
For showing $(iv) \Rightarrow (ii)$, let $\sum_k x_{n_k}$ be any subseries of $\sum_n x_n$. Let $\varepsilon > 0$. Then by the assumption of (iv), there is a positive integer N such that $|\sum_{i \in A} x_i| < \varepsilon$ whenever A is a finite subset of $\mathbb{Z}_+$ with $\min A > N$. Choose K such that $n_k > N$ for all $k \ge K$. This implies that $|x_{n_{k+1}} + \cdots + x_{n_{k+p}}| < \varepsilon$ for all $k > K$ and for all $p = 1, 2, ...$, so the series $\sum_k x_{n_k}$ is convergent.
For $(ii) \Rightarrow (iv)$, assume that (iv) does not hold. As in the proof of $(i) \Rightarrow (iv)$, there is $\varepsilon > 0$ and there is a subsequence (x_{n_i}) such that $|\sum_{n_i \in A_k} x_{n_i}| \ge \varepsilon$, thus, the subseries $\sum_i x_{n_i}$ is divergent.
For $(iv) \Rightarrow (i)$, let π be any bijection on $\mathbb{Z}_+$. Let $\varepsilon > 0$ and let N be given as in (iv). Take i_0 such that $\pi(i) > N$ for all $i \ge i_0$. This implies that $|\sum_{i_1 \le i \le i_2} x_{\pi(i)}| < \varepsilon$ for all $i_0 \le i_1 < i_2$. Thus, (i) holds.
The proof is finished. □

Remark 5.8. *Notice that from the proof of Theorem 5.9, we see that the Theorem does still hold if the series $\sum x_n$ is taken in $\mathbb{R}^N$.*

Corollary 5.9. *Let (x_n) be a sequence of real numbers. Then $\sum_n x_n$ is absolutely convergent if and only if it is unconditionally convergent.*

Proof. Part ($\Rightarrow$) has been shown in Theorem 5.5.
For ($\Leftarrow$), assume that $\sum_n x_n$ is unconditionally convergent. For each n, let $\varepsilon_n := \pm 1$ such that $|x_n| = \varepsilon_n x_n$. Then by Theorem 5.9 $(i) \Leftrightarrow (iii)$, the series $\sum |x_n| = \sum \varepsilon_n x_n$ is convergent as desired. The proof is finished. □

6. 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 6.1. *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 6.2. *In Lemma 6.9(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 6.3. *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 6.4. *Notice that by Lemma 6.9, 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 6.5. *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 6.6. *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 6.7. *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 6.8. *Let $f(x) = \sum x^n/n^2$. Then by using the same argument of Example 6.7, we have $r = 1$. On the other hand, it is known that $f(\pm 1)$ both are convergent. So dom $f = [-1, 1]$.*

Lemma 6.9. *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 6.1 at once. □

Remark 6.10. *Note that the Example 6.7 shows us that f may not converge uniformly on $(-r, r)$. In fact let f be defined as in Example 6.7. 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 6.11. *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 6.12. *With the notation as above if $0 < r \leq \infty$, then $f \in C^\infty(-r, r)$. Moreover, the k-derivatives $f^{(k)}(x) = \sum_{n \geq k} a_k n(n-1)(n-2) \cdots\cdots (n-k+1) x^{n-k}$ for all $x \in (-r, r)$.*

Proof. Fix $c \in (-r, r)$. By Lemma 6.9, 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 $\lim_n |n a_n|^{1/n} = \lim_n |a_n|^{1/n}$. 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 6.13. *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 6.9 f converges uniformly on $[0, x]$. Since each term $a_n t^n$ is continuous, the result follows. $\square$

Theorem 6.14. (**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

$$\begin{aligned} s_{n+p}(x) - s_n(x) &= a_{n+1}x^{n+1} + a_{n+2}x^{n+1} + a_{n+3}x^{n+1} + \cdots\cdots\cdots + a_{n+p}x^{n+1} \\ &+ a_{n+2}(x^{n+2} - x^{n+1}) + a_{n+3}(x^{n+2} - x^{n+1}) + \cdots\cdots\cdots + a_{n+p}(x^{n+2} - x^{n+1}) \\ &+ a_{n+3}(x^{n+3} - x^{n+2}) + \cdots\cdots\cdots + a_{n+p}(x^{n+3} - x^{n+2}) \\ &\vdots \\ &+ a_{n+p}(x^{n+p} - x^{n+p-1}). \end{aligned} \tag{6.1}$$

Since $x \in [0, 1]$, $|x^{n+k+1} - x^{n+k}| = x^{n+k} - x^{n+k+1}$. So the Eq.6.1 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. $\square$

Remark 6.15. *In Remark 6.10, 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.*

7. Real analytic functions

Proposition 7.1. *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. $\square$

Definition 7.2. *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 7.3.

(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 6.12, we have seen that $f\in C^{\infty}(a,b)$ and*

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

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

(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 $k=0,1,2\ldots.$ 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^{∞}.*

Lemma 7.4. *Let $(a_{jk})_{j,k\in\mathbb{N}}$ be a set of real numbers. Assume that $\sum_{k=0}^{\infty}\sum_{j=0}^{k}|a_{jk}|<\infty$. Then $L:=\sum_{k=0}^{\infty}\sum_{j=0}^{k}a_{jk}$ exists and $L=\sum_{j=0}^{\infty}\sum_{j\geq k}a_{jk}$.*

Proof. Note that $L:=\sum_{k=0}^{\infty}\sum_{j=0}^{k}a_{jk}$ exists due to the Cauchy theorem. Put $b_k:=\sum_{j=0}^{k}|a_{jk}|$. Then for any $\varepsilon>0$, there is $K_1\in\mathbb{N}$ such that

$$|L-\sum_{k=0}^{K}\sum_{j=0}^{k}a_{jk}|<\varepsilon \text{ for all } \quad K>K_1; \qquad \text{and} \qquad \sum_{k>K_1}\sum_{j=0}^{k}|a_{jk}|<\varepsilon.$$

Let $J_1=K_1$. Then for $J>J_1$, we have

$$\begin{aligned}|L-\sum_{j=0}^{J}\sum_{k\geq j}a_{jk}| &\leq |L-\sum_{j=0}^{J_1}\sum_{k\geq j}a_{jk}|+|\sum_{J_1<j\leq J}\sum_{k\geq j}a_{jk}|\\ &\leq |L-\sum_{0\leq j\leq J_1}\sum_{j\leq k\leq K_1}a_{jk}|+|\sum_{0\leq j\leq J_1}\sum_{k>K_1}a_{jk}|+|\sum_{J_1<j\leq J}\sum_{k\geq j}a_{jk}|\\ &<3\varepsilon.\end{aligned}$$

□

Proposition 7.5. *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 $c = 0$ and hence, $I = (-r, r)$. Now fix $z \in I$ and 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 $|z| + |x-z| \in I$ for all $x \in (z-\delta, z+\delta)$ and thus, $\sum_{k=0}^{\infty} |a_k|(|z| + |x-z|)^k < \infty$. Lemma 7.4 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 7.6. *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

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

whenever $|x| < 1$.
Consequently, $(1+x)^\alpha$ is analytic on $(-1,1)$.

Proof. Considering the formal power series

$$F(x) := \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k.$$

The ratio test implies that the radius of the series $F(x)$ is $r = 1$. Hence, the series $F(x)$ is convergent in $(-1,1)$. In particular, $F(x)$ is analytic on $(-1,1)$ by Proposition 7.5. We are going to show that $F(x) = (1+x)^\alpha$ for all $x \in (-1,1)$. Notice that we have the following equation.

$$(1+x)F'(x) = \alpha F(x) \quad \text{for all } x \in (-1,1). \tag{7.1}$$

To see this note that we have

$$F'(x) = \sum_{n=1}^{\infty} n\binom{\alpha}{n} x^{n-1} = \sum_{j=0}^{\infty} (j+1)\binom{\alpha}{j+1} x^j.$$

This is obtained by the following direct calculation.

$$\begin{aligned}(1+x)F'(x) &= \sum_{j=0}^{\infty}(j+1)\binom{\alpha}{j+1}x^j + \sum_{n=1}^{\infty} n\binom{\alpha}{n}x^n \\ & \alpha + \sum_{j=1}^{\infty}\{(j+1)\binom{\alpha}{j+1} + j\binom{\alpha}{j}\}x^j \\ &= \alpha + \sum_{j=1}^{\infty}\alpha\binom{\alpha}{j}x^j.\end{aligned}$$

Thus, the Eq 7.1 holds. From this we have $F(x) \neq 0$ for all $x \in (-1,1)$. To see this, if $F(c) = 0$ for some $c \in (-1,1)$, then $F'(c) = 0$ by Eq 7.1. Differentiating the Eq 7.1, we get $F^{(2)}(c) = 0$. To repeat the same step, we have $F^{(n)}(c) = 0$ for all $n = 0, 1, 2....$ Notice that since F is real analytic on $(-1,1)$, $F(x) = \sum_{n=0}^{\infty} \frac{F^{(n)}(c)}{n!}(x-c)^n$ in some open subinterval J of $(-1,1)$ that contains c and so $F \equiv 0$ on J. From this if we put the set $L(c) := \{-1 < \alpha < c : F(t) = 0; \forall t \in (\alpha, c]\}$, then $L(c) \neq \emptyset$. Hence, $a := \inf L(c)$ exists and so $-1 \leq a$. First we notice that $a \in L(c)$, that is $F|(a,c] \equiv 0$. Next we want to show that $a = -1$. If not, assume $-1 < a$. Since $F(a) = \lim_{t\to a+} F(t)$, we have $F(a) = 0$. As the reason above, there is an open subinterval J_1 of $(-1,1)$ containing a satisfying $F|J_1 \equiv 0$ and so, there is a point $-1 < a_1 < a$ such that $F|(a_1, a] \equiv 0$. This gives $a_1 \in L(c)$ and so, $a \leq a_1$ that contradicts to $a_1 < a$. Therefore, we have $F|(-1,c] \equiv 0$. Similarly, one can also obtain $F|[c,1) \equiv 0$. Hence, $F \equiv 0$ on $(-1,1)$. It is absurd. This and Eq 7.1 give

$$\int_0^x \frac{F'(t)}{F(t)}\,dt = \int_0^x \frac{\alpha}{1+t}\,dt$$

for all $x \in (-1,1)$. This implies that $F(x) = (1+x)^{\alpha}$ as desired. □

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