NOTE ON MATH2060B: ELEMENTARY ANALYSIS II (2020-21)

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 16, 2021.

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

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

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 $L:=\lim\limits_{x\to c}\dfrac{f'(x)}{g'(x)}$ exists, then so does $\lim\limits_{x\to c}\dfrac{f(x)}{g(x)}$, moreover, we have $L=\lim\limits_{x\to c}\dfrac{f(x)}{g(x)}$.*

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 $L:=\lim\limits_{x\to c}\dfrac{f'(x)}{g'(x)}$ exists, then $\lim\limits_{x\to c+}\dfrac{f(x)}{g(x)}$ exists and is equal to L.

Similarly, we also have $\lim\limits_{x\to c-}\dfrac{f(x)}{g(x)}=L$. 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 $f'(c)=\lim\limits_{t\to 0-}\dfrac{f(c+t)-f(c-t)}{2t}$. 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 (⇒): assume that f is a convex function. Fix a point $c \in (a,b)$. f is convex, so we have $f(c) = f(\frac{1}{2}(c+t) + \frac{1}{2}(c-t)) \le \frac{1}{2}f(c+t) + \frac{1}{2}f(c-t)$ 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 $f^{(2)}(c) \ge 0$.
For (⇐), assume that $f^{(2)}(x) \ge 0$ 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

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

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 $p \ge 1$.*

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 $|x_2-z|\leq M\cdot|z-x_1|^2\leq(M\delta)\cdot|x_1-z|<\delta$. 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. □

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.

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

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

Example 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 8.3, 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 $\omega_i(f,P)\le\varepsilon$ and thus,

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

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

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

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

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

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

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

The proof is complete. $\square$

Proposition 2.14. *We have the following assertions.*

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

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

Proof. For Part (i), it is clear that we have the inequality $U(f,P)\le U(g,P)$ for any partition P. So, we have $\int_a^b f = \overline{\int_a^b} f \le \overline{\int_a^b} g = \int_a^b g$.
For Part (ii), the integrability of $|f|$ follows immediately from Theorem 8.3 and the simple inequality $||f|(x') - |f|(x'')| \le |f(x') - f(x'')|$ for all $x',x''\in[a,b]$. Thus, we have $U(|f|,P) - L(|f|,P) \le$

$U(f,P) - L(f,P)$ for any partition P on $[a,b]$.
Finally, since we have $-f \le |f| \le f$, by Part (i), we have $|\int_a^b f| \le \int_a^b |f|$ at once. □

Proposition 2.15. *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*

$$\int_a^b f = \int_a^c f + \int_c^b f. \tag{2.4}$$

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

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

for any partition P_1 on $[a,c]$ and P_2 on $[c,b]$ with $P = P_1 \cup P_2$.
From this, we can show the sufficient condition at once.
For showing the necessary condition, since $f \in R[a,b]$, for any $\varepsilon > 0$, there is a partition Q on $[a,b]$ such that $U(f,Q) - L(f,Q) < \varepsilon$ by Theorem 8.3. 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 2.4 above. Notice that for any partition P_1 on $[a,c]$ and P_2 on $[c,b]$, we have

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

So, we have $\int_a^c f + \int_c^b f \le \int_a^b f$. Then the inverse inequality can be obtained at once by considering the function $-f$. Then the resulted is obtained by using Theorem 8.3. □

Proposition 2.16. *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 8.3, 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 2.17. *In the proof of Proposition 2.16, 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 2.18. (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

$$\int_a^b f(x)g(x)dx = f(\xi)\int_a^b g(x)dx. \tag{2.5}$$

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

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

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

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

for all $x \in [a,b]$. From this and Proposition 2.16 above, we have

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

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

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

Therefore, there is a point $\xi \in [x_1, x_2] \subseteq [a,b]$ so that the Equation 2.5 holds by using the Intermediate Value Theorem for the function f. Thus, it remains to show that such element ξ can be chosen in (a,b).
Let $a \le x_1 < x_2 \le b$ be as above.
If x_1 and x_2 can be found so that $a < x_1 < x_2 < b$, then the result is proved immediately since $\xi \in [x_1, x_2] \subset (a,b)$ in this case.
Now suppose that x_1 or x_2 does not exist in (a,b), i.e., $m = f(a) < f(x)$ for all $x \in (a,b]$ or $f(x) < f(b) = M$ for all $x \in [a,b)$.
Claim 1: If $f(a) < f(x)$ for all $x \in (a,b]$, then $\int_a^b fg > f(a)\int_a^b g$ and hence, $\xi \in (a, x_2] \subseteq (a,b]$.
For showing **Claim1**, put $h(x) := f(x) - f(a)$ for $x \in [a,b]$. Then h is continuous on $[a,b]$ and $h > 0$ on $(a,b]$. This implies that $\int_c^d h > 0$ for any subinterval $[c,d] \subseteq [a,b]$. (**Why**?)
On the other hand, since $\underline{\int_a^b} g = \int_a^b g > 0$, there is a partition $P : a = x_0 < \cdots < x_n = b$ so that $L(g,P) > 0$. This implies that $m_k(g,P) > 0$ for some sub-interval $[x_{k-1}, x_k]$. Therefore, we have

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

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

Now if $f \in R[a,b]$, then by Proposition 2.15, we can define a function $F : [a,b] \to \mathbb{R}$ by

$$F(c) = \begin{cases} 0 & \text{if } c = a \\ \int_a^c f & \text{if } a < c \le b. \end{cases} \tag{2.6}$$

Theorem 2.19. 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. 2.6 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 $\lim\limits_{t\to 0+} \frac{1}{t}(F(c+t) - F(c)) = f(c)$. Similarly, we have $\lim\limits_{t\to 0-} \frac{1}{t}(F(c+t) - F(c)) = f(c)$. So, we have $F'(c) = f(c)$ as desired. The proof is finished. □

Definition 2.20. *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 $A = \lim\limits_{\|P\|\to 0} R(f,P,\{\xi_i\})$, if for any $\varepsilon > 0$, there is $\delta > 0$ such that

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

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

Proposition 2.21. *Let f be a function defined on $[a,b]$. If the limit $\lim\limits_{\|P\|\to 0} R(f,P,\{\xi_i\}) = A$ exists, then f is automatically bounded.*

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

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

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

Lemma 2.22. *$f \in R[a,b]$ 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 8.3.
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

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

and

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

The proof is finished. □

Theorem 2.23. *$f\in R[a,b]$ if and only if the Riemann sum $R(f,P,\{\xi_i\})$ is convergent. In this case, $R(f,P,\{\xi_i\})$ converges to $\int_a^b f(x)dx$ as $\|P\|\to 0$.*

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

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

and

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

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

Now let $\varepsilon>0$. Lemma 2.22 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 2.21.
Now fix a partition P with $\|P\|<\delta$. Then for each $[x_{i-1},x_i]$, choose $\xi_i\in[x_{i-1},x_i]$ such that $M_i(f,P)-\varepsilon\leq f(\xi_i)$. This implies that we have

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

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

$$\overline{\int_a^b} f(x)dx\leq U(f,P)\leq A+\varepsilon(1+b-a). \tag{2.7}$$

By considering $-f$, note that the Riemann sum of $-f$ will converge to $-A$. The inequality 8.1 will imply that for any $\varepsilon > 0$, there is a partition P such that

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

The proof is complete. □

Theorem 2.24. *Let $f \in R[c,d]$ and let $\phi : [a,b] \longrightarrow [c,d]$ be a strictly increasing C^1 function with $f(a) = c$ and $f(b) = d$.*
Then $f \circ \phi \in R[a,b]$, moreover, we have

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

Proof. Let $A = \int_c^d f(x)dx$. By using Theorem 2.23, 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 2.22 and Theorem 2.23, there is $\delta_1 > 0$ such that

$$|A - \sum f(\eta_k)\triangle x_k| < \varepsilon \tag{2.8}$$

and

$$\sum \omega_k(f, P)\triangle x_k < \varepsilon \tag{2.9}$$

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

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

This yields that

$$|\triangle x_k - \phi'(\xi_k)\triangle t_k| < \varepsilon\Delta t_k \tag{2.10}$$

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

$$\begin{aligned} |A - \sum f(\phi(\xi_k))\phi'(\xi_k)\triangle t_k| &\leq |A - \sum f(\phi(\xi_k^*))\phi'(\xi_k^*)\triangle t_k| \\ &\quad + |\sum f(\phi(\xi_k^*))\phi'(\xi_k^*)\triangle t_k - \sum f(\phi(\xi_k^*))\phi'(\xi_k)\triangle t_k| \\ &\quad + |\sum f(\phi(\xi_k^*))\phi'(\xi_k)\triangle t_k - \sum f(\phi(\xi_k))\phi'(\xi_k)\triangle t_k| \end{aligned} \tag{2.11}$$

Notice that inequality 8.2 implies that

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

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

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

where $|f(x)| \le M$ for all $x \in [c,d]$.
On the other hand, by using inequality 8.4 we have

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

for all k. This, together with inequality 8.3 imply that

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

Finally by inequality 8.5, we have

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

The proof is complete. □

3. Improper Riemann Integrals

Definition 3.1. *Let $-\infty < a < b < \infty$.*

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

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

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

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

Example 3.2. *Define (*formally*) an improper integral $\Gamma(s)$ (called the Γ-function) as follows:*

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

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

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

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

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

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

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

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

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. □

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

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

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 9.1 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.9.3 implies that

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

for all $m \geq N$ and for all $x \in (a, b)$ with $x \neq u$.
Taking $m \to \infty$ in Eq.9.4, we have

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

Hence we have

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

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

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

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.9.5 give

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

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*

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

On the other hand, $f_n \to 0$ uniformly on $(-1,1)$. In fact, if $f_n'(1/n) = 0$ for all $n = 1,2,..,$ then f_n attains the maximal value $f_n(1/n) = \frac{1}{2n}$ at $x = 1/n$ for each $n = 1,...$ and hence, $f_n \to 0$ uniformly on $(-1,1)$.
So Propositions 9.1 and 9.2 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 $z := \lim\limits_k x_{n_k} \in A$. Since $g_{n_k}(z) \downarrow 0$ as $k \to \infty$. So, there is a positive integer K such that $0 \leq g_{n_K}(z) < \varepsilon/2$. Since g_{n_K} is continuous at z and $\lim\limits_i x_{n_i} = z$, we have $\lim\limits_i g_{n_K}(x_{n_i}) = g_{n_K}(z)$. So, we can choose i large enough such that $i > K$

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

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

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

We can now conclude that

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

The proof is complete. □

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 11.9(ii), notice that if $f(c)$ is convergent, it does not imply f converges uniformly on $[-c, c]$ in general.*
For example, $f(x) := 1 + \sum_{n=1}^{\infty} \dfrac{x^n}{n}$. 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 11.9, 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 11.7, 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 11.1 at once. □

Remark 6.10. *Note that the Example 11.7 shows us that f may not converge uniformly on $(-r, r)$. In fact let f be defined as in Example 11.7. 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 11.9, one can choose $0 < \eta < r$ such that $c \in (-\eta,\eta)$ and f converges uniformly on $(-\eta,\eta)$.
It needs to show that the k-derivatives $f^{(k)}(c)$ exists for all $k \geq 0$. Consider the case $k = 1$ first.
If we consider the series $\sum_{n=0}^\infty (a_n x^n)' = \sum_{n=1}^\infty n a_n x^{n-1}$, then it also has the same radius r because $\lim_n |na_n|^{1/n} = \lim_n |a_n|^{1/n}$. This implies that the series $\sum_{n=1}^\infty n a_n x^{n-1}$ converges uniformly on $(-\eta,\eta)$. Therefore, the restriction $f|(-\eta,\eta)$ is differentiable. In particular, $f'(c)$ exists and $f'(c) = \sum_{n=1}^\infty n a_n c^{n-1}$.
So the result can be shown inductively on k. □

Proposition 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 11.9 f converges uniformly on $[0,x]$. Since each term $a_n t^n$ is continuous, the result follows. □

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

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

Since $x \in [0,1]$, $|x^{n+k+1} - x^{n+k}| = x^{n+k} - x^{n+k+1}$. So the Eq.11.1 implies that
$|s_{n+p}(x)-s_n(x)| \leq \varepsilon(x_{n+1}+(x^{n+1}-x^{n+2})+(x^{n+2}-x^{n+3})+\cdots+(x^{n+p-1}-x^{n+p})) = \varepsilon(2x^{n+1}-x^{n+p}) \leq 2\varepsilon.$
So f converges uniformly on $[0,1]$ as desired.
Finally for the general case, we consider $g(x) := f(rx) = \sum_n a_n r^n x^n$. Note that $\lim_n |a_n r^n|^{1/n} = 1$ and $g(1) = f(r)$. Then by the case above,, we have shown that

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

The proof is finished. □

Remark 6.15. *In Remark 11.10, 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. □

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 11.12, we have seen that $f \in C^\infty(a,b)$ and

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

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

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

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

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

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

Proposition 7.4. *Suppose that $f(x) := \sum_{k=0}^\infty a_k(x-c)^k$ is convergent on some open interval I centered at c, that is $I = (c-r, c+r)$ for some $r > 0$. Then f is analytic on I.*

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

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

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

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

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

Example 7.5. *Let $\alpha \in \mathbb{R}$. Recall that $(1+x)^\alpha$ is defined by $e^{\alpha \ln(1+x)}$ for $x > -1$.*
Now for each $k \in \mathbb{N}$, put

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

Then

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

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

Proof. Notice that $f^{(k)}(x) = \alpha(\alpha-1)\cdot\cdots\cdot(\alpha-k+1)(1+x)^{\alpha-k}$ for $|x| < 1$.
Fix $|x| < 1$. Then by Proposition 12.1, for each positive integer n we have

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

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

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

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

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

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

On the other hand, note that we always have $0 < 1-\eta_n < 1+\eta_n x$ for all n because $x > -1$. Thus, we can now conclude that $R_n(x) \to 0$ as $|x| < 1$. The proof is finished. Finally the last assertion follows from Proposition 12.4 at once. The proof is complete. □