Math 230B Lecture Notes

Lance Remigio

July 15, 2025

Chapter 1

Week 1

1.1 Lecture 1

1.1.1 Topics

- The derivative
- Continuity and Differentiability
- Differentiability Rules

Definition (Differentiability). (*) Let $I \subseteq \mathbb{R}$ be an interval, $f : I \to \mathbb{R}$, $c \in I$. We say f is **differentiable** at c if

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

exists (that is, it equals a real number).

(*) In this case, the quantity $\lim_{x \to c} \frac{f(x)-f(c)}{x-c}$ is called the derivative of f at c and is denoted by

$$f'(c), \frac{df}{dx}(c), \left.\frac{df}{dx}\right|_{x=c}$$

(*) If $f : I \to \mathbb{R}$ is differentiable at every point $c \in I$, we say f is differentiable (on I).

Remark. The following are equivalent characterizations of the differentiability:

$$\begin{aligned} f'(c) = L &\Longleftrightarrow \lim_{x \to c} \frac{f(x) - f(c)}{x - c} = L \\ &\Longleftrightarrow \forall \varepsilon > 0 \; \exists \delta > 0 \text{ such that if } 0 < |x - c| < \delta \text{ then } \left| \frac{f(x) - f(c)}{x - c} - L \right| < \varepsilon \\ &\Longleftrightarrow \forall \varepsilon > 0 \; \exists \delta > 0 \text{ such that if } 0 < |h| < \delta \text{ then } \left| \frac{f(c+h) - f(c)}{h} - L \right| < \varepsilon \\ &\Longleftrightarrow \lim_{h \to 0} \frac{f(c+h) - f(c)}{h} = L \end{aligned}$$

Theorem (Differentiability Implies Continuous). Let $I \subseteq \mathbb{R}$, $c \in I$, and $f : I \to \mathbb{R}$ is differentiable at c. Then f is continuous at c.

Proof. It suffices to show that $\lim_{x\to c} f(x) = f(c)$. Note that

$$\begin{aligned}\lim_{x\to c}(f(x)-f(c)) &= \lim_{x\to c}\left[\frac{f(x)-f(c)}{x-c}\right](x-c)\\ &= \left[\lim_{x\to c}\frac{f(x)-f(c)}{x-c}\right]\left[\lim_{x\to c}(x-c)\right]\\ &= (f'(c))(0)\\ &= 0.\end{aligned}$$

So, we have

$$\begin{aligned}\lim_{x\to c} f(x) &= \lim_{x\to c}[f(x)-f(c)+f(c)]\\ &= \lim_{x\to c}(f(x)-f(c)) + \lim_{x\to c} f(c)\\ &= 0 + \lim_{x\to c} f(c)\\ &= 0 + f(c)\\ &= f(c).\end{aligned}$$

■

Corollary. If $f : I \to \mathbb{R}$ is NOT continuous at $c \in I$, then f is NOT differentiable at c.

Example. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} x^2 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q} \end{cases}.$$

(i) Prove that f is continuous at 0.
Proof. Our goal is to show that

$$\forall \varepsilon > 0\ \exists \delta > 0 \text{ such that if } |x| < \delta \text{ then } |f(x) - f(c)| < \varepsilon.$$

Let $\varepsilon > 0$ be given. Note that if $x \notin \mathbb{Q}$,

$$|f(x)| = |0| < \varepsilon.$$

Otherwise, we have $|f(x)| = |x^2| = |x|^2$. IN this case, we claim that $\delta = \sqrt{\varepsilon}$ will work. Indeed, if $|x| < \delta$, then we have

$$|f(x)| = |x|^2 < (\sqrt{\varepsilon})^2 = \varepsilon.$$

■

(ii) Prove f is discontinuous at all $x \neq 0$.
Proof. Let $c \neq 0$. Our goal is to show that f is discontinuous at c. By the sequential criterion for continuity, it suffices to find a sequence (a_n) such that $a_n \to c$ but $f(a_n) \not\to f(c)$. We will consider two cases; that is, we could either have $c \notin \mathbb{Q}$ or $c \in \mathbb{Q}$.

Suppose $c \notin \mathbb{Q}$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, there exists a sequence of rational numbers (r_n) such that $r_n \to c$. Note that $f(r_n) = r_n^2 \to c^2 \neq 0$, but $f(c) = 0$. Clearly, $f(r_n) \not\to f(c)$ and so f must be discontinuous at c.

Suppose $c \in \mathbb{Q}$. Since the set of irrational numbers is also dense in $\mathbb{R}$, we can find a sequence (s_n) such that $s_n \to c$. Note that $f(s_n) = 0$, but $f(c) = c^2 \neq 0$. Thus, $f(s_n) \not\to f(c)$. Therefore, f must be discontinuous at c. ■

(iii) Prove that f is nondifferentiable at all $x \neq 0$.
Proof. Let $c \neq 0$. Since f is discontinuous at c, we can conclude that f is not differentiable at c. ■

(iv) Prove that $f'(0) = 0$.

Proof. We need to show

$$\lim_{x \to c} \frac{f(x) - f(0)}{x - 0} = \frac{f(x)}{x} = 0.$$

■

Theorem (Algebraic Differentiability Theorem). Assume that $f : I \to \mathbb{R}$ and $g : I \to \mathbb{R}$ are differentiable at $c \in I$ where (I is an interval on $\mathbb{R}$). Then

(i) For all $k \in \mathbb{R}$, kf is differentiable at c, and

$$(kf)'(c) = kf'(c)$$

(ii) $f + g$ is differentiable at c, and

$$(f + gk)'(c) = f'(c) + g'(c)$$

(iii) fg is differentiable at c, and

$$(fg)'(c) = f'(c)g(c) + f(c)g'(c)$$

(iv) $\frac{f}{g}$ is differentiable at c provided that $g(c) \neq 0$. Then

$$\left(\frac{f}{g}\right)'(c) = \frac{f'(c)g(c) - f(c)g'(c)}{[g(c)]^2}.$$

Theorem (Chain Rule). Let $I_1 \subseteq \mathbb{R}$ and $I_2 \subseteq \mathbb{R}$ be two intervals, $f : I_1 \to \mathbb{R}$ and $g : I_2 \to \mathbb{R}$ be two functions, $f(I_1) \subseteq I_2$, f is differentiable at $c \in I_1$ and g is differentiable at $f(c) \in I_2$. Then the function $g \circ f : I_1 \to \mathbb{R}$ is differentiable at $c \in I_1$ and

$$(g \circ f)'(c) = g'(f(c)) \cdot f'(c).$$

First, we will prove the theorem incorrectly and then show give three criterion to prove the theorem correctly.

Proof. Observe that

$$\begin{aligned}
\lim_{x \to c} \frac{(g \circ f)(x) - (g \circ f)(c)}{x - c} &= \lim_{x \to c} \frac{g(f(x)) - g(f(c))}{x - c} \\
&= \lim_{x \to c} \left[\frac{g(f(x)) - g(f(c))}{f(x) - f(c)} \cdot \frac{f(x) - f(c)}{x - c}\right] \\
&= \underbrace{\left[\lim_{x \to c} \frac{g(f(x)) - g(f(c))}{f(x) - f(c)}\right]}_{g'(f(c))} \underbrace{\left[\lim_{x \to c} \frac{f(x) - f(c)}{x - c}\right]}_{f'(c)}
\end{aligned}$$

■

What is the problem with this proof? By the definition of a limit of a function, when you take $\lim\limits_{x \to c}$, it is guaranteed that $x - c \neq 0$; however, for x close to c (as x approaches to c), $f(x) - f(c)$ might be zero, so dividing by $f(x) - f(c)$ is not legitimate. The following proof fixes this issue by introducing a new function $d(f(x))$ which is defined by

(i) $d(f(x))\dfrac{g(f(x)) - g(f(c))}{f(x) - f(c)}$ when $f(x) \neq f(c)$

(ii) $d(f(x))$ is defined even when $f(x) = f(c)$

(iii) $d(f(x)) \cdot \dfrac{f(x)-f(c)}{x-c} = \dfrac{g(f(x))-g(f(c))}{x-c}$ for all $x \in I_\kappa$ where $x \neq c$.

Proof. Let $d : I_2 \to \mathbb{R}$ be defined by

$$d(y) = \begin{cases} \frac{g(y)-g(f(c))}{y-f(c)} & \text{if } y \neq f(c) \\ \lim\limits_{y\to f(c)} \frac{g(y)-g(f(c))}{y-f(c)} = g'(f(c)) & \text{if } y = f(c) \end{cases}.$$

Note that this function satisfies the requirements in (i) and (ii) outlined above. We make the following observations:

(1) d is continuous at $f(c)$. Indeed, we can see that

$$\lim_{y\to f(c)} d(y) = \lim_{y\to f(c)} \frac{g(y)-g(f(c))}{y-f(c)} = d(f(c)).$$

(2) For all $x \in I_1$ and $x \neq c$, we have

$$d(f(x)) \cdot \frac{f(x)-f(c)}{x-c} = \frac{g(f(x))-g(f(c))}{x-c}. \tag{*}$$

We will show that this holds by considering two cases; either $f(x) \neq f(c)$ or $f(x) = f(c)$. If $f(x) \neq f(c)$, then

$$\text{LHS} = d(f(x)) \cdot \frac{f(x)-f(c)}{x-c} = \frac{g(f(x))-g(f(c))}{f(x)-f(c)} \cdot \frac{f(x)-f(c)}{x-c} = \frac{g(f(x))-g(f(c))}{x-c} = \text{RHS}.$$

Now, suppose $f(x) = f(c)$. Then we have

$$\begin{aligned} \text{LHS} &= d(f(x)) \cdot \frac{f(x)-f(c)}{x-c} = d(f(c)) \cdot \frac{f(x)-f(c)}{x-c} = g'(f(c)) \cdot \frac{0}{x-c} = 0 \\ \text{RHS} &= \frac{g(f(x))-g(f(c))}{x-c} = \frac{g(f(c))-g(f(c))}{x-c} = \frac{0}{x-c} = 0. \end{aligned}$$

Thus, we see that the left hand side equals the right hand side of (*).

Now, note that since f is continuous at c and d is continuous at $f(c)$, their composition $d \circ f$ is continuous at c and so,

$$\lim_{x\to c}(d\circ f)(x) = (d\circ f)(c).$$

$$\begin{aligned} &\lim_{x\to c} \frac{(g\circ f)(x)-(g\circ f)(c)}{x-c} = \lim_{x\to c} \frac{g(f(x))-g(f(c))}{x-c} \\ &= \left[\lim_{x\to c}(d\circ f)(x)\right] \cdot \left[\lim_{x\to c} \frac{f(x)-f(c)}{x-c}\right] \\ &= [(d\circ f)(c)] \cdot f'(c) \\ &= [d(f(c))] \cdot f'(c) \\ &= g'(f(c)) \cdot f'(c). \end{aligned}$$

■

1.2 Lecture 2-4

1.2.1 Topics

(1) Local Maxima and minima

(2) Interior Extremum Theorem (Theorem 5.8)

(3) Darboux's Theorem (Theorem 5.12)

(4) Some observations

(5) Rolle's Theorem

(6) Mean Value Theorem

Theorem (Interior Extremum Theorem)**.** Let $I \subseteq \mathbb{R}$ be an interval and $f : I \to \mathbb{R}$. Suppose c is an interior point of I and f is differentiable at c. Then

(1) If f has a local max at c, then $f'(c) = 0$;

(2) If f has a local min at c, then $f'(c) = 0$.

Before we prove this theorem, we will first go over an important lemma that is used in the main proof.

Lemma. Suppose $\lim_{x\to c} g(x)$ and $\lim_{x\to c} h(x)$ both exist.

(1) If there exists $\delta > 0$ such that $h(x) \leq g(x)$ for all $x \in (c-\delta, c)$, then $\lim_{x\to c} h(x) \leq \lim_{x\to c} g(x)$.

(2) If there exists $\delta > 0$ such that $h(x) \leq g(x)$ for all $x \in (c, c+\delta)$, then $\lim_{x\to c} h(x) \leq \lim_{x\to c} g(x)$.

Proof. Here we will prove (1). The proof of (2) is analogous. Let (a_n) be a sequence in $(c-\delta, c)$ such that $a_n \to c$. By the Sequential Criterion for limits of functions, we have $a_n \to c$ implies $\lim_{n\to\infty} g(a_n) = \lim_{x\to c} g(x)$ and $\lim_{n\to\infty\infty} h(a_n) = \lim_{x\to c} h(x)$. Also, note that from the Order Limit Theorem for sequences, we can see that

$$\begin{aligned} \forall n \; a_n \in (c-\delta, c) &\Longrightarrow \forall n \; h(a_n) \leq g(a_n) \\ &\Longrightarrow \lim_{n\to\infty} h(a_n) \leq \lim_{n\to\infty} g(a_n). \end{aligned}$$

Hence, we can see from these two observations that

$$\lim_{x\to c} h(x) \leq \lim_{x\to c} g(x).$$

■

1.2.2 Proof of the Interior Extremum Theorem

Proof. Here we will prove (1). Suppose f has a local max at c. Then

- If f has a local max at c, then there exists $\delta_1 > 0$ such that for all $x \in (c-\delta_1, c+\delta_1) \cap I$ $f(x) \leq f(c)$.
- If c is an interior point of I, then there exists $\delta_2 > 0$ such that $(c-\delta_2, c+\delta_2) \subseteq I$. So, if we let $\delta = \min\{\delta_1, \delta_2\}$, then
$$\forall x \in (c-\delta, c+\delta) \quad f(x) \leq f(c).$$

We have

(I) For all $x \in (c-\delta, c)$, we see that $x - c < 0$ and $f(x) \leq f(c)$ implies that

$$\frac{f(x) - f(c)}{x - c} \geq 0.$$

By the Order Limit Theorem for functions, we have

$$\lim_{x\to c} \frac{f(x) - f(c)}{x - c} \geq \lim_{x\to c} 0 \Longrightarrow f'(c) \geq 0.$$

(II) For all $x \in (c, c+\delta)$. Since $x - c > 0$ and $f(x) \leq f(c)$, we have

$$\frac{f(x) - f(c)}{x - c} \leq 0.$$

Using the Order Limit Theorem again, we have

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c} \leq \lim_{x \to c} 0 \Longrightarrow f'(c) \leq 0.$$

From (I) and (II), we can see that $f'(c) \leq 0$ and $f'(c) \geq 0$. Thus, $f'(c) = 0$. ■

Theorem (Darboux's Theorem)**.** Let $f : [a, b] \to \mathbb{R}$ be differentiable and $f'(a) < f'(b)$ (or $f'(b) < f'(a)$). Let $\alpha \in \mathbb{R}$ be such that $f'(a) < \alpha < f'(b)$ (or $f'(b) < \alpha < f'(a)$). Then there exists $c \in (a, b)$ such that $f'(c) = \alpha$.

Proof. Let $g : [a, b] \to \mathbb{R}$ be defined by $g(x) = f(x) - \alpha x$. It follows from the Algebraic Differentiability Theorem that g is differentiable on $[a, b]$ and so it is continuous on $[a, b]$. It suffices to show that there is a point $c \in (a, b)$ such that $g'(c) = 0$. To this end, it is enough to show that there exists a point $c \in (a, b)$ at which g has a local min. Since g is continuous on $[a, b]$ and $[a, b]$ is compact, g attains its minimum on $[a, b]$. Let $\hat{c}$ be a point at which g attains a minimum. IN what follows we will show that $\hat{c} \in (a, b)$ and so it can be used as the same c that we were looking for. Note that

$$g'(a) = f'(a) - \alpha < 0$$
$$g'(b) = f'(b) - \alpha > 0.$$

We make two claims; namely, $\hat{c} \neq a$ ($\hat{c}$ cannot be the left-endpoint of $[a, b]$) and $\hat{c} \neq b$ (the right-endpoint of $[a, b]$). Suppose for contradiction that $\hat{c} = a$. Then for all $x \in [a, b]$ $g(x) \geq g(a)$. So, for all $x \in (a, b)$, we have $g(x) - g(a) \geq 0$ and $x - a > 0$. Thus, for all $x \in (a, b)$

$$\frac{g(x) - g(a)}{x - a} \geq 0.$$

Thus,

$$\lim_{x \to a} \geq \frac{g(x) - g(a)}{x - a} \geq 0 \Longrightarrow \lim_{x \to a} 0.$$

That is, $g'(a) \geq 0$. This contradicts the fact that $g'(a) < 0$ and this proves the first claim. A similar argument shows that $g'(b) \leq 0$, contradicting the fact that $g'(b) > 0$. ■

Theorem (Rolle's Theorem)**.** Let $f : [a, b] \to \mathbb{R}$ be continuous, let f be differentiable on (a, b), and $f(a) = f(b)$. Then there exists a point $c \in (a, b)$ such that $f'(c) = 0$.

Proof. It suffices to show that there exists a point $c \in (a, b)$ for which f has a local max or a local min. Since f is continuous on $[a, b]$ and $[a, b]$ is compact, we see that f must attain its max and min on $[a, b]$. Thus, we may consider two cases; namely, both $\max f$ and $\min f$ occur at the endpoints and or they occur in the interior of $[a, b]$. Considering the first case, we see that from the fact that $f(a) = f(b)$ that

$$\max_{a \leq x \leq b} f(x) = \min_{a \leq x \leq b} f(x).$$

Hence, f must be constant on $[a, b]$ and so,

$$\forall x \in [a, b] \quad f'(x) = 0.$$

In this case, we can choose c to be any point we like in (a, b).

Now, assume the second case. Using the Interior Extremum Theorem, we have $f'(c) = 0$ and this concludes the proof. ■

Theorem (Mean Value Theorem)**.** Let $f : [a, b] \to \mathbb{R}$ be a continuous function and f is differentiable in (a, b). Then there exists a $c \in (a, b)$ such that

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

Proof. Let $g : [a, b] \to \mathbb{R}$ be defined by

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

By the Algebraic Continuity Theorem, g must be continuous on $[a, b]$. Also, the Algebraic Differentiability Theorem implies that g is differentiable on (a, b). Then we have

$$g(a) = f(a) - \frac{f(b) - f(a)}{b - a} = \frac{bf(a) - af(a) + af(a)}{b - a} = \frac{bf(a) - af(b)}{b - a}$$

and

$$g(b) = f(b) - \frac{f(b) - f(a)}{b - a}b = \frac{bf(a) - af(b) - bf(b) + bf(a)}{b - a} = \frac{bf(a) - af(b)}{b - a}.$$

Using Rolle's Theorem, it follows that

$$\exists c \in (a, b) \;\; g'(c) = 0.$$

Note that

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}.$$

Thus,

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

■

Theorem (Generalized Mean Value Theorem)**.** Let $f : [a, b] \to \mathbb{R}$ and $g : [a, b] \to \mathbb{R}$ are continuous functions and suppose f and g are differentiable in (a, b). Then there exists a point $c \in (a, b)$ such that

$$f'(c)[g(b) - g(a)] = g'(c)[f(b) - f(a)].$$

Proof. See proof in hw1. ■

Remark. Note that, if g' is never zero in (a, b), then we may rewrite the claim of GMVT as follows:

$$\exists c \in (a, b) \;\text{ such that } \frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Corollary (Corollary 1)**.** Let $I \subseteq \mathbb{R}$ be an interval, $f : I \to \mathbb{R}$ be diffferentiable, and $f'(x) = 0$ for all $x \in I$. Then f is a constant function on I, that is, there exists $k \in \mathbb{R}$ such that for all $x \in I$, $f(x) = k$.

Proof. Let $x, y \in I$ with x, y. It suffices to show that $f(x) = f(y)$. To this end, we will apply the MVT to f on the interval $[x, y]$. By the Mean Value Theorem, there exists a $c \in (x, y)$ such that

$$f'(c) = \frac{f(y) - f(x)}{y - x}.$$

Since $f'(x) = 0$ for all $x \in I$, we have $f'(c) = 0$ and so

$$0 = \frac{f(y) - f(x)}{y - x} \implies 0 = f(y) - f(x) \implies f(y) = f(x).$$

■

Corollary. Let $I \subseteq \mathbb{R}$ is an interval, $f : I \to \mathbb{R}$ and $g : I \to \mathbb{R}$ are differentiable, and $f'(x) = g'(x)$ for all $x \in I$. Then there exists $k \in \mathbb{R}$ such that for all $x \in I$, $f(x) = g(x) + k$.

Proof. Let $h = f - g$. We have

$$\forall x \in I \quad h'(x) = (f - g)'(x) = f'(x) - g'(x) = 0.$$

Using Corollary 1, we have

$$\exists k \in \mathbb{R} \text{ such that } \forall x \in I \; h(x) = k.$$

Hence, we have

$$\exists k \in \mathbb{R} \text{ such that } \forall x \in I \; f(x) - g(x) = k$$

and so $f(x) = g(x) + k$ as desired. ■

Theorem (Parts (a) and (c)). Let $I \subseteq \mathbb{R}$ is an interval and suppose $f : I \to \mathbb{R}$ is differentiable. Then

(1) f is increasing on I if and only if $\forall c \in I \; f'(c) \geq 0$.

(2) f is decreasing on I if and only if $c \in I \; f'(c) \leq 0$.

Proof. Here we will prove (1). The proof of (2) is similar. ($\implies$) Let f be increasing on I. Let $c \in I$. Note that for all $x \in I$ with $x \neq c$. Then we have

$$\frac{f(x) - f(c)}{x - c} \geq 0.$$

Indeed, we can see that if $x > c$, then $x - c > 0$ and $f(x) \geq f(c)$. Thus, we have that

$$\frac{f(x) - f(c)}{x - c} \geq 0.$$

If $x < c$, then $x - c < 0$ and $f(x) \leq f(c)$. Then

$$\frac{f(x) - f(c)}{x - c} \geq 0.$$

It follows from the Order Limit Theorem that

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c} \geq \lim_{x \to c} 0 = 0.$$

Hence, $f'(c) \geq 0$ as desired.

$\Longleftarrow$ Suppose for all $c \in I$, $f'(c) \geq 0$. To show that f is increasing on I, we need to show that for all $x_1, x_2 \in I$ with $x_1 < x_2$. It suffices to show that $f(x_1) \leq f(x_2)$. To this end, we will apply the Mean Value Theorem to the function f on $[x_1, x_2]$. That is, there exists a $c \in (x_1, x_2)$ such that

$$f'(c) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Thus,

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1).$$

Since $f'(c) \geq 0$ and $x_2 - x_1 \geq 0$, we can see that

$$f(x_2) \geq f(x_1)$$

as desired. ■

Theorem (L'Hopital's Rule). Let $I \subseteq \mathbb{R}$ be an interval, $a \in I$, and $f : I \to \mathbb{R}$ and $g : I \to \mathbb{R}$ are continuous functions. Suppose f and g are differentiable at all points in $I \setminus \{a\}$ with $f(a) = g(a) = 0$ and

$$\lim_{x \to a} \frac{f'(x)}{g'(x)} = L \in \mathbb{R}.$$

Then

$$\lim_{x \to a} \frac{f(x)}{g(x)} = L.$$

Proof. Our goal is to show that

$$\forall \varepsilon > 0 \ \exists \delta > 0 \text{ such that if } 0 < |x - c| < \delta \text{ with } (x \in I) \text{ then } \Big| \frac{f(x)}{g(x)} - L \Big| < \varepsilon. \qquad (*)$$

Let $\varepsilon > 0$ be given. Since, by assumption,

$$\lim_{x \to a} \frac{f'(x)}{g'(x)} = L,$$

for the given $\varepsilon > 0$, there exists $\hat{\delta} > 0$ such that if $0 < |x - a| < \hat{\delta}$ (with $x \in I$, then

$$\Big| \frac{f'(x)}{g'(x)} - L \Big| < \varepsilon.$$

We claim that this $\hat{\delta}$ works as the δ we were looking for. Indeed, if we let $\delta = \hat{\delta}$, then (*) will hold. The reason is as follows: suppose $x \in I$ and $0 < |x - a| < \delta$. In what follows, we will show that

$$\Big| \frac{f(x)}{g(x)} - L \Big| < \varepsilon.$$

We may consider the following cases:

(1) $(x > a)$ (that is, $x \in I$ and $x \in (a, a + \hat{\delta})$). We apply the Generalized Mean Value Theorem to f and g on the interval $[a, x]$. That is,

$$\exists c \in (a, x) \text{ such that } \frac{f'(c)}{g'(c)} = \frac{f(x) - f(a)}{g(x) - g(a)}.$$

Since $f(a) = g(a) = 0$, we can conclude that

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

It follows that

$$\Big| \frac{f(x)}{g(x)} - L \Big| = \Big| \frac{f'(c)}{g'(c)} - L \Big| < \varepsilon.$$

Note that the latter inequality is true because $0 < |c - a| \leq |x - a| < \hat{\delta}$.

(2) $(x < a)$ (that is, $x \in I$ and $x \in (a - \hat{\delta}, a)$) We apply the Generalized Mean Value Theorem to f and g on $[x, a]$. That is, there exists $c \in (x, a)$ such that

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

and so,

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

It follows that

$$\Big| \frac{f(x)}{g(x)} - L \Big| = \Big| \frac{f'(c)}{g'(c)} - L \Big| < \varepsilon.$$

Note that the latter inequality is true because $0 < |c - a| \leq |x - a| < \hat{\delta}$.

■

Chapter 2

Week 2

2.1 Topics

- A Useful Theorem (Another consequence of the GMVT)
- Taylor Polynomials
- Taylor's Theorem with Lagrange remainder

Corollary (A corollary of GMVT)**.** Let $I \subseteq \mathbb{R}$ be an open interval, $x_0 \in I$, and $n \in \mathbb{N} \cup \{0\}$, $f : I \to \mathbb{R}$ has $n+1$ derivatives, and $f^{(k)}(x_0) = 0$ for all $0 \leq k \leq n$. Then for each point $x \neq x_0$ in the interval I, there exists a point c strictly between x and x_0 such that

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

2.1.1 Observation 1

Let $k \in \mathbb{N}$ and let $x_0 \in \mathbb{R}$ fixed. Then

$$\begin{aligned}
\frac{d}{dx}[(x-x_0)^k] &= k(x-x_0)^{k-1} \\
\frac{d^2}{dx^2}[(x-x_0)^k] &= \frac{d}{dx}[k(x-x_0)^{k-1}] = k(k-1)(x-x_0)^{k-2} \\
&\vdots \\
\frac{d^k}{dx^k}[(x-x_0)^k] &= k(k-1)\cdots(1)(x-x_0)^{k-k} = k!
\end{aligned}$$

where

$$\frac{d^j}{dx^j}[(x-x_j)^k] = k(k-1)\cdots(k-(j-1))(x-x_0)^{k-j}.$$

2.1.2 Observation 2

$$\frac{d^j}{dx^j}[(x-x_0)^k] = \begin{cases} k(k-1)\cdots(k-j+1)(x-x_0)^{k-j} & \text{if } j < k \\ k! & \text{if } j = k \\ 0 & \text{if } j > k \end{cases}$$

and

$$\frac{d^j}{dx^j}[(x-x_0)^k]\Big|_{x=x_0} = \begin{cases} 0 & \text{if } j < k \\ k! & \text{if } j = k \\ 0 & \text{if} j > k. \end{cases}$$

With these two observations, we will now prove the claim made in the corollary above.

Proof. Here we will prove the claim for the case where $x > x_0$. The proof for $x < x_0$ is completely analogous. Let $g : I \to \mathbb{R}$ be defined by $g(t) = (t - x_0)^{n+1}$. Note that

$$g^{(k)}(x_0) = 0 \; \forall 0 \le k \le n \qquad \text{(for } t \neq 0 \; g^{(k)(t)} \neq 0\text{)}$$
$$g^{(n+1)}(t) = (n+1)! \; \forall t \in I.$$

Now, we apply GMVT to f and g on the interval $[x_0, x]$: Using the GMVT, we can find $x_1 \in (x_0, x)$ such that

$$\frac{f'(x_1)}{g'(x_1)} = \frac{f(x) - f(x_0)}{g(x) - g(x_0)}$$

and so

$$\frac{f'(x_1)}{g'(x_1)} = \frac{f(x)}{g(x)}. \tag{I}$$

Next, we apply GMVT to find f' and g' on the interval $[x_0, x_1]$: Using the GMVT, we can find an $x_2 \in (x_0, x_1)$ such that

$$\frac{f"(x_2)}{g"(x_2)} = \frac{f'(x_1) - f'(x_0)}{g'(x_1) - g'(x_0)} \underbrace{=}_{f'(x_0)=0, g'(x_0)=0} \frac{f'(x_1)}{g'(x_1)} \underbrace{=}_{(I)} = \frac{f(x)}{g(x)}.$$

Continuing in this manner, we will obtain $x_{n+1} \in (x_0, x)$ such that

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

and so,

$$\frac{f^{(n+1)}(x_{n+1})}{(n+1)!} = \frac{f(x)}{(x - x_0)^{n+1}}.$$

Thus,

$$f(x) = \frac{f^{(n+1)}(x_{n+1})}{(n+1)!}(x - x_0)^{n+1}$$

■

What are the nicest functions that we know? Which functions are easiest to work with? Polynomials! Another question that we would like to answer is:

Given a function f, is it possible to find a "good" approximation for f among polynomials?

To understand the situation a bit better, we break the questions asked above in the following three questions:

- Does there exist a polynomial $p(x)$ such that

$$p(x_0) = f(x_0)$$
$$p'(x_0) = f'(x_0)$$
$$\vdots$$
$$p^{(n)}(x_0) = f^{(n)}(x_0).$$

  That is, does there exist a polynomial $p(x)$ that agrees with f to order n at x_0.

- If such a polynomial $p(x)$ exists, how can we find it?
- Suppose we use $p(x)$ as an approximation of $f(x)$ near x_0; how good is this approximation? What can be said about the error?

Remark. (i) Number of equations to be satisfied by $p(x)$ is $n + 1$.

(ii) Also note that a polynomial in $\P_n$ can be represented as $c_0 + c_2x + \cdots + c_nx^n$ (we have $n + 1$ coefficients).

From (i) and (ii), it seems reasonable to expect that we might be able to find a polynomial $p(x)$

in $\P_n$ that satisfies all the equations.

2.1.3 Answers to Q1 and Q2

There is a **unique** polynomial in $\P_n$ that agrees with $f(x)$ to order n at $x_0 \in I$ in the sense that

$$\begin{aligned} p(x_0) &= f(x_0) \\ &\vdots \\ p^{(n)}(x_0) &= f^{(n)}(x_0). \end{aligned}$$

This polynomial can be denoted by

$$T_{n,x_0}(x)$$

which is the n**th Taylor Polynomial of** f **centered at** x_0. Moreover, we have

$$\begin{aligned} T_{n,x_0}(x) &= f(x_0)'_f(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \cdots + \frac{f^{(n)(x_0)}}{n!}(x-x_0)^n \\ &= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k. \end{aligned}$$

2.1.4 Proof of our Observation

Proof. Let $p(x)$ be a general polynomial of degree at most n:

$$p(x) = c_0 + c_1(x-x_0) + \cdots + c_n(x-x_0)^n.$$

Our goal is to show that if $p^{(\ell)}(x_0) = f^{(\ell)}(x_0)$, then

$$p(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k \forall a \le \ell \le n.$$

Note that $p(x_0) = c_0$. Also, for $1 \le \ell \le n$, we have

$$\begin{aligned} p^{(\ell)} &= \frac{d^\ell}{dx^\ell}\left[c_0 + \sum_{k=1}^{n} c_k(x-x_0)^k\right] \\ &= \frac{d^\ell}{dx^\ell}\left[\sum_{k=1}^{n} c_k(x-x_0)^k\right] \\ &= \sum_{k=1}^{n} c_k \frac{d^\ell}{dx^\ell}[(x-x_0)^k]. \end{aligned}$$

Hence, we have

$$p^{(\ell)}(x_0) = \sum_{k=1}^{n} c_k \frac{d^\ell}{dx^\ell}[(x-x_0)^k]\Big|_{x=x_0} = c_\ell \ell! \ .$$

Therefore,

$$\forall 1 \le \ell \le n \ \ p^{(\ell)}(x_0) = c_\ell \ell! \ .$$

We see that p agrees with f to order n at x_0 if and only if $p(x_0) = f(x_0)$ and $p^{(\ell)}(x_0) = f^{(\ell)(x_0)}$ for all $1 \le \ell \le n$. This is true if and only if

$$\begin{aligned} c_0 &= f(x_0) \\ \ell! c_\ell = f^{(\ell)}(x_0) \forall 1 \le \ell \le n. \end{aligned}$$

Furthermore, this is true if and only if

$$\begin{aligned} c_0 &= f(x_0) \\ c_\ell &= \frac{f^{(\ell)(x_0)}}{\ell!} \forall 1 \le \ell \le n. \end{aligned}$$

That is,

$$\begin{aligned} p(x) &= \sum_{k=0}^{n} c_k (x - x_0)^k = c_0 + \sum_{k=1}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k \\ &= f(x_0) + \sum_{k=1}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k \\ &= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k. \end{aligned}$$

■

Theorem (Taylor's Theorem with Lagrange Remainder). Let $I \subseteq \mathbb{R}$ be an open interval, $x_0 \in I$, and $n \in \mathbb{N} \cup \{0\}$. Suppose $f : I \to \mathbb{R}$ has $n+1$ derivatives. Then for each point $x \neq x_0$ in I, there is a point c strictly between x and x_0 such that

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

where

$$\begin{aligned} T_{n,x_0}(x) &= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{n!}(x - x_0)^k \\ R_{n,x_0}(x) &= \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}. \end{aligned}$$

Remark.
- Note that clearly the equality above holds at $x = x_0$ too (for any value of c).
- Recall that for any fixed number R

$$\lim_{n \to \infty} \frac{R^{n+1}}{(n+1)!} = 0.$$

However, $f^{(n+1)}(c)$ may become very large.

2.1.5 Proof of Taylor's Theorem

Proof. Let $R_{n,x_0}(x) = f(x) - T_{n,x_0}(x)$. Our goal is to show that

$$R_{n,x_0}(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}$$

for some c between x and x_0. Note that

(i) By assumption, f contains $n+1$ derivatives and T_{n,x_0} is a polynomial which also contains $n+1$ derivatives. Also, we have

$$R_{n,x_0} = f - T_{n,x_0}.$$

Thus, R_{n,x_0} must have $n+1$ derivatives.

(ii) For all $0 \leq k \leq n$,

$$R^{(k)}_{n,x_0}(x_0) = f^{(k)}(x_0) - T^{(k)}_{n,x_0}(x_0) = 0.$$

Using (i), (ii), and corollary 5, we can see that for each $x \neq x_0$, we have

$$R_{n,x_0}(x) = \frac{R^{(n+1)}_{n,x_0}(c)}{(n+1)!}(x - x_0)^{n+1} \tag{I}$$

for some c strictly between x and x_0. Now, note that

$$R_{n,x_0}^{(n+1)}(c) = f^{(n+1)}(c) - T_{n,x_0}^{(n+1)}(c) = f^{(n+1)}(c). \tag{II}$$

Using (I) and (II), we have

$$R_{n,x_0} = \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}.$$

■

Chapter 3

Week 3

Chapter 4

Week 4

4.1 Lecture 6

4.2 Lecture 6

4.2.1 Topics

(1) The definition of Riemann-Stieltjes integral

(2) Refinement of partitions

Definition (Almost Disjoint Intervals). We say that two intervals I and J are **almost disjoint** if either $I \cap J$ is empty or $I \cap J$ has exactly one point.

Definition (Partition). A partition P of an interval $[a, b]$ is a finite set of points in $[a, b]$ that includes both a and b. We always list the points of a partition $P = \{x_0, x_1, x_2, \ldots, x_n\}$ in an increasing order; so,

$$a = x_0 < x_1 < \cdots < x_n = b.$$

Remark. A partition of P of an interval $[a, b]$ is a finite collection of almost disjoint (nonempty) compact intervals whose union is $[a, b]$:

$$P = I_1, I_2, \ldots, I_n$$

where

$$I_1 = [x_0, x_1], \quad I_2 = [x_1, x_2], \quad \cdots \quad I_n = [x_{n-1}, x_n].$$

Again, we denote $x_0 = a$ and $x_n = b$.

Definition (Lower Sum, Upper Sum). Let $f : [a, b] \to \mathbb{R}$ be bounded, $\alpha : [a, b] \to \mathbb{R}$ be increasing, and $P = \{x_0, x_2, \ldots, x_n\}$ be a partition of $[a, b]$. Let $\Delta\alpha_k = \alpha(x_k) - \alpha(x_{k-1})$.

(i) The **Lower Riemann-Stieltjes Sum** of f with respect to the integrator α for the partition P is defined by

$$L(f, \alpha, P) = \sum_{k=1}^{n} m_k(\alpha(x_k) - \alpha(x_{k-1})) = \sum_{k=1}^{n} m_k \Delta\alpha_k.$$

(ii) The upper **Riemann-Stieltjes sum** of f with respect to the integrator α for the partition P

is defined by

$$U(f,\alpha,P)=\sum_{k=1}^{n}M_k(\alpha(x_k)-\alpha(x_{k-1}))=\sum_{k=1}^{n}M_k\Delta\alpha_k.$$

Definition (Upper R.S Integral, Lower R.s Integral)**.** Let $f:[a,b]\to\mathbb{R}$ be bounded, $\alpha:[a,b]\to\mathbb{R}$ be increasing. Then

(i) The **Upper R.S integral** of f with respect to α (on $[a,b]$) is defined by

$$U(f,\alpha)=\inf_{P\in\Pi}U(f,\alpha,P).$$

Note that the set $\{U(f,\alpha,P):P\in\Pi\}$ is bounded below by $m(\alpha(b)-\alpha(a))$. So the infimum above is a real number.

(ii) The **Lower R.S Integral** of f with respect to α (on $[a,b]$) is defined by

$$L(f,\alpha)=\sup_{P\in\Pi}L(f,\alpha,P).$$

Note that the set $\{L(f,\alpha,P):P\in\Pi\}$ the lower sums is bounded above by $M(\alpha(b)-\alpha(a))$. So, the supremum above is a real number.

Definition (Riemann-Stieltjes integrable functions)**.** Let $\alpha:[a,b]\to\mathbb{R}$ be an increasing function. A function $f:[a,b]\to\mathbb{R}$ is said to be **Riemann-Stieltjes integrable** (on $[a,b]$) if

(i) f is bounded

(ii) $L(f,\alpha)=U(f,\alpha)$.

In this case, the R.S integral of f with respect to α, denoted by

$$\int_a^b f\ d\alpha \quad\text{or}\quad \int_a^b f(x)\ d\alpha(x) \quad\text{or}\quad \int_{[a,b]} f\ d\alpha$$

is the common value of $L(f,\alpha)$ and $U(f,\alpha)$. That is,

$$\int_a^b f\ d\alpha=L(f,\alpha)=U(f,\alpha).$$

4.3 Lecture 8-9-10

Theorem (Rudin 6.4)**.** Let $f:[a,b]\to\mathbb{R}$ be bounded, $\alpha:[a,b]\to\mathbb{R}$ is increasing, P is a partition of $[a,b]$, and Q is a refinement of P. Then

(1) $L(f,\alpha,P)\le L(f,\alpha,Q)$

(2) $U(f,\alpha,P)\ge U(f,\alpha,Q)$

Proof. Here we will prove (1). The proof of (2) is completely analagous. We proceed via induction on $\ell=\operatorname{card}(Q\setminus P)$ (the number of points in $Q\setminus P$). Let $P=\{x_0,x_1,\dots,x_n\}$.

If $\ell=0$, then $P\subseteq Q$ and $\operatorname{card}Q=\operatorname{card}P$ implies that $P=Q$. Thus, $L(f,\alpha,P)=L(f,\alpha,Q)$.

If $\ell=1$, then Q has exactly one extra point. Let's call this point z. So, $\{z\}=Q\setminus P$. Note that

$z \in [a,b]$ and P is a partition of $[a,b]$. Hence, there exists $1 \leq i \leq n$ such that $z \in (x_{i-1}, x_i)$. Let

$$m_i' = \inf_{x \in [x_{i-1}, z]} f(x)$$
$$m_i'' = \inf_{x \in [z, x_i]} f(x)$$

Recall that if $A \subseteq B$, then $\inf A \geq \inf B$. Hence, $m_i' \geq m_i$ and $m_i'' \geq m_i$. We have

$$\begin{aligned}
L(f,\alpha,P) &= \sum_{k=1}^{n} m_k(\alpha(x_k)) \\
&= \Big[\sum_{k \neq i} m_k(\alpha(x_k) - \alpha(x_{k-1}))\Big] + m_i(\alpha(x_i) - \alpha(z) + \alpha(z) - \alpha(x_{i-1})) \\
&= \Big[\sum_{k \neq i} m_k(\alpha(x_k) - \alpha(x_{k-1}))\Big] + m_i(\alpha(z) - \alpha(x_{i-1})) + m_i(\alpha(x_i) - \alpha(z)) \\
&\leq \Big[\sum_{k \neq i} m_k(\alpha(x_k) - \alpha(x_{k-1}))\Big] + m_i'(\alpha(z) - \alpha(x_{i-1})) + m_i''(\alpha(x_i) - \alpha(z)) \\
&= L(f,\alpha,Q).
\end{aligned}$$

So, we have $L(f,\alpha,P) \leq L(f,\alpha,Q)$.

Now, suppose the claim is true for $\ell = r \geq 1$. Our goal is to show that the claim holds for $\ell = r+1$. Suppose $\mathrm{card}(Q \setminus P) = r+1$. Let

$$Q \setminus P = \{z_1, z_2, \ldots, z_r, z_{r+1}\}.$$

Let $\hat{Q} = P \cup \{z_1, z_2, \ldots, z_r\}$. We have

$$L(f,\alpha,P) \leq L(f,\alpha,\hat{Q}) \leq L(f,\alpha,Q)$$

where the first inequality holds due to our induction hypothesis and the second inequality holds because $Q \setminus \hat{Q}$ contains only one point. So, we have

$$L(f,\alpha,P) \leq L(f,\alpha,Q).$$

■

> **Theorem.** Let $f : [a,b] \to \mathbb{R}$ be a bounded function, $\alpha : [a,b] \to \mathbb{R}$ is increasing. Let P_1 and P_2 are any two partition of $[a,b]$. Then
> $$L(f,\alpha,P_1) \leq U(f,\alpha,P_2).$$

Proof. Let $Q = P_1 \cup P_2$ be the common refinement of P_1 and P_2. Applying the previous theorem, we can see that $P_1 \subseteq P_1 \cup P_2$ and $P_2 \subseteq P_1 \cup P_2$ implies

$$L(f,\alpha,P_1) \leq L(f,\alpha,Q) \leq U(f,\alpha,Q) \leq U(f,\alpha,P_2)$$

■

For the following theorem, we will use the lemma below.

> **Lemma.** Suppose A and B are nonempty subsets of $\mathbb{R}$. If
> $$\forall a \in A \ \forall b \in B \quad a \leq b$$
> then $\sup A \leq \inf B$.

Theorem (Rudin 6.5)**.** Let $f : [a, b] \to \mathbb{R}$ be a bounded function and $\alpha : [a, b] \to \mathbb{R}$ is an increasing function. Then $L(f, \alpha) \leq U(f, \alpha)$.

Proof. Let $A = \{L(f, \alpha, P) : P \in \Pi\}$ and $B = \{U(f, \alpha, P) : P \in \Pi\}$. Using the lemma above and Theorem 2, we can see that for all $a \in A$ and for all $b \in B$, it follows that $\sup A \leq \inf B$; that is, $L(f, \alpha) \leq U(f, \alpha)$. ■

Theorem (Cauchy Criterion for Riemann-Stieltjes Integrability Rudin 6.6)**.** Let $f : [a, b] \to \mathbb{R}$ be a bounded function, $\alpha : [a, b] \to \mathbb{R}$ be an increasing function. Then

$$f \in R_\alpha[a, b] \iff \forall \varepsilon > 0 \ \exists P_\varepsilon \in \Pi[a, b] \text{ such that } U(f, \alpha, P_\varepsilon) - L(f, \alpha, P_\varepsilon) < \varepsilon.$$

Proof. ($\Longleftarrow$) Our goal is to show that $L(f, \alpha) = U(f, \alpha)$. Note that $L(f, \alpha) \leq U(f, \alpha)$ implies $U(f, \alpha) - L(f, \alpha) \geq 0$. Hence, it suffices to show that for all $\varepsilon > 0$,

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

Let $\varepsilon > 0$ be given. By assumption, there exists $P_\varepsilon \in \Pi$ such that

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

We have

$$\begin{aligned} U(f, \alpha) &= \inf_{P \in \Pi} U(f, \alpha, P) \leq U(f, \alpha, P_\varepsilon) \\ L(f, \alpha) &= \sup_{P \in \Pi} L(f, \alpha, P) \geq L(f, \alpha, P_\varepsilon) \end{aligned}$$

Using Rudin 6.5, we can see that

$$L(f, \alpha, P_\varepsilon) \leq L(f, \alpha) \leq U(f, \alpha) \leq U(f, \alpha, P_\varepsilon).$$

So, the interval $[L(f, \alpha), U(f, \alpha)]$ is contained in the interval $[L(f, \alpha, P_\varepsilon), U(f, \alpha, P_\varepsilon)]$. Thus,

$$U(f, \alpha) - L(f, \alpha) \leq U(f, \alpha, P_\varepsilon) - L(f, \alpha, P_\varepsilon) < \varepsilon$$

as desired.

($\Longrightarrow$) Our goal is to show that for any $\varepsilon > 0$, there exists a partition $P_\varepsilon \in \Pi$ such that

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

Note that

$$\begin{aligned} U(f, \alpha) &= \inf_{P \in \Pi} U(f, \alpha, P) \Longrightarrow \exists P_1 \in \Pi \text{ such that } U(f, \alpha, P_1) < U(f, \alpha) + \frac{\varepsilon}{2} \\ L(f, \alpha) &= \sup_{P \in \Pi} L(f, \alpha, P) \Longrightarrow \exists P_2 \in \Pi \text{ such that } L(f, \alpha) - \frac{\varepsilon}{2} < L(f, \alpha, P_2) \end{aligned}$$

Let $P_\varepsilon = P_1 \cup P_2$ (we claim that this partition can be used as the one that we were looking for).

$$L(f, \alpha) - \frac{\varepsilon}{2} < L(f, \alpha, P_2) \leq L(f, \alpha, P_\varepsilon) \leq U(f, \alpha, P_\varepsilon) \leq U(f, \alpha, P_1) < U(f, \alpha) + \frac{\varepsilon}{2}.$$

Thus, we have

$$\begin{aligned} U(f, \alpha, P_\varepsilon) - L(f, \alpha, P_\varepsilon) &< \left[\left(U(f, \alpha) + \frac{\varepsilon}{2}\right) - \left(L(f, \alpha) - \frac{\varepsilon}{2}\right)\right] \\ &= U(f, \alpha) - L(f, \alpha) + \varepsilon \\ &= 0 + \varepsilon = \varepsilon \end{aligned}$$

as desired. ■

Theorem (Rudin 6.7). Let $f : [a, b] \to \mathbb{R}$ be a bounded function, $\alpha : [a, b] \to \mathbb{R}$ is an increasing function, fix $\varepsilon > 0$, $P = \{x_0, x_1, \ldots, x_n\}$ is a partition of $[a, b]$, and

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

Then

(1) If Q is any refinement of P, then $U(f, \alpha, Q) - L(f, \alpha, Q) < \varepsilon$.

(2) If for every $1 \leq k \leq n$, t_k and s_k are arbitrary points in $[x_{k-1}, x_k]$, then

$$\sum_{k=1}^{n} |f(s_k) - f(t_k)|\Delta\alpha_k < \varepsilon.$$

(3) If $f \in R_\alpha[a, b]$ and for each $1 \leq k \leq n$, s_k is a point in $[x_{k-1}, x_k]$, then

$$\Big|\sum_{k=1}^{n} f(s_k)\Delta\alpha_k - \int_a^b f \; d\alpha\Big| < \varepsilon.$$

Proof. (1) We have

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

Therefore,

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

(2) For each $1 \leq k \leq n$, we have

$$\begin{aligned} &m_k \leq f(s_k) \leq M_k \\ &m_k \leq f(t_k) \leq M_k \Longrightarrow -M_k \leq -f(t_k) \leq -m_k. \end{aligned}$$

So, we have

$$m_k - M_k \leq f(s_k) - f(t_k) \leq M_k - m_k.$$

That is,

$$-(M_k - m_k) \leq f(s_k) - f(t_k) \leq M_k - m_k.$$

Therefore,

$$|f(s_k) - f(t_k)| \leq M_k - m_k.$$

Hence, we have

$$\sum_{k=1}^{n} |f(s_k) - f(t_k)|\Delta\alpha_k \leq \sum_{k=1}^{n} (M_k - m_k)\Delta\alpha_k = U(f, \alpha, P) - L(f, \alpha, P) < \varepsilon.$$

(3) For all $1 \leq k \leq n$, we have

$$m_k \leq f(s_k) \leq M_k.$$

So,

$$\sum_{k=1}^{n} m_k\Delta\alpha_k \leq \sum_{k=1}^{n} f(s_k)\Delta\alpha_k \leq \sum_{k=1}^{n} M_k\Delta\alpha_k.$$

Therefore,

$$L(f, \alpha, P) \leq \sum_{k=1}^{n} f(s_k) \leq U(f, \alpha, P) \tag{I}$$

Also, note that

$$L(f, \alpha, P) \leq \int_a^b f \; d\alpha \leq U(f, \alpha, P). \tag{II}$$

Hence,

$$\Big|\sum_{k=1}^{n} f(s_k)\Delta\alpha_k - \int_a^b f \; d\alpha\Big| \leq U(f, \alpha, P) - L(f, \alpha, P) < \varepsilon$$

as desired.

■

Lemma. Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Let $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a, b]$. Then

$$\forall 1 \leq k \leq n \quad \sup_{s,t \in [x_{k-1}, x_k]} |f(s) - f(t)| = M_k - m_k.$$

Proof. Let $k \in \{1, 2, \ldots, n\}$. We need to show

(1) $\forall s, t \in [x_{k-1}, x_k]$ $|f(s) - f(t)| \leq M_k - m_k$.

(2) $\forall \varepsilon > 0$, $\exists \hat{s}\hat{t} \in [x_{k-1}, x_k]$ such that $M_k - m_k - \varepsilon < |f(\hat{s}) - f(\hat{t})|$.

Note that we have already shown (1) in our discussion of Theorem 6.7.

Let $\varepsilon > 0$ be given. Then we have

$$\begin{aligned} m_k &= \inf_{t \in [x_{k-1}, x_k]} f(t) \Longrightarrow \hat{t} \in [x_{k-1}, x_k] \text{ such that } f(\hat{t}) < m_k + \frac{\varepsilon}{2} \\ M_k &= \sup_{t \in [x_{k-1}, x_k]} f(t) \Longrightarrow \hat{s} \in [x_{k-1}, x_k] \text{ such that } M_k - \frac{\varepsilon}{2} < f(\hat{s}). \end{aligned}$$

Adding the inequalities above, we get

$$M_k - m_k - \varepsilon < f(\hat{s}) - f(\hat{t}) \leq |f(\hat{s}) - f(\hat{t})|.$$

■

Theorem (Rudin 6.8)**.** Let $f : [a, b] \to \mathbb{R}$ be a continuous function and $\alpha : [a, b] \to \mathbb{R}$ is an increasing function. Then $f \in R_\alpha[a, b]$.

Proof. Since $f : [a, b] \to \mathbb{R}$ is a continuous function and $[a, b]$ is compact, it follows from the Extreme Value Theorem that f is bounded on $[a, b]$. Now, according to the Cauchy Criterion for Riemann-Stieltjes integrability, it suffices to show that

$$\forall \varepsilon > 0 \ \exists P \in \Pi \text{ such that } U(f, \alpha, P) - L(f, \alpha, P) < \varepsilon. \tag{*}$$

Let $\varepsilon > 0$ be given. By the same reasoning that showed f is bounded on $[a, b]$, it follows that f is uniformly continuous on $[a, b]$. For the given ε, there exists a $\delta > 0$ such that for all $s, t \in [a, b]$:

$$\text{if } |s - t| < \delta \text{ then } |f(s) - f(t)| < \frac{\varepsilon}{2[\alpha(b) - \alpha(a) + 1]}.$$

Let $P = \{x_0, x_1, \ldots, x_n\}$ be any partition of $[a, b]$ such that $\|P\| < \delta$. We claim (*) holds for such a partition. Indeed, for all $k \in \{1, 2, \ldots, n\}$ and for all $s, t \in [x_{k-1}, x_k]$, if $|s - t| < \delta$, then

$$|f(s) - f(t)| < \frac{\varepsilon}{2[\alpha(b) - \alpha(a) + 1]}.$$

Hence,

$$\sup_{s,t \in [x_{k-1}, x_k]} |f(s) - f(t)| \leq \frac{\varepsilon}{2[\alpha(b) - \alpha(a) + 1]}.$$

Thus,

$$M_k - m_k \leq \frac{\varepsilon}{2[\alpha(b) - \alpha(a) + 1]}.$$

Therefore,

$$
\begin{aligned}
U(f,\alpha,P) - L(f,\alpha,P) &= \sum_{k=1}^{n}(M_k - m_k)\Delta\alpha_k \\
&\le \sum_{k=1}^{n} \frac{\varepsilon}{2[\alpha(b)-\alpha(a)+1]}\Delta\alpha_k \\
&= \frac{\varepsilon}{2[\alpha(b)-\alpha(a)+1]}\sum_{k=1}^{n}\Delta\alpha_k \\
&= \frac{\varepsilon}{2[\alpha(b)-\alpha(a)+1]}\cdot[\alpha(b)-\alpha(a)] \\
&\le \frac{\varepsilon}{2} \\
&< \varepsilon
\end{aligned}
$$

as desired.

■

Lemma. Let $\alpha : [a,b] \to \mathbb{R}$ be an increasing and continuous function and $\alpha(a) < \alpha(b)$. Then for each $n \in \mathbb{N}$, there exists a partition $P = \{x_0, x_1, \dots, x_n\}$ such that

$$\forall 1 \le k \le n \quad \Delta\alpha_k = \alpha(x_k) - \alpha(x_{k-1}) = \frac{\alpha(b)-\alpha(a)}{n}.$$

Proof. Let $n \in \mathbb{N}$. Divide the interval $[\alpha(a), \alpha(b)]$ into n subintervals of equal length: $\frac{\alpha(b)-\alpha(a)}{n}$. For each $1 \le k \le n$, we have $y_k \in (\alpha(a), \alpha(b))$. Hence, the Intermediate Value Theorem implies that

$$\exists x_k \in (a,b) \text{ such that } y_k = \alpha(x_k).$$

Since α is increasing, we have

$$a = x_0 < x_1 < x_2 < \cdots < x_n = b.$$

This tells us that $P = \{x_0, x_1, \dots, x_n\}$ will be partition of $[a,b]$ such that

$$\forall 1 \le k \le n \quad \Delta\alpha_k = \alpha(x_k) - \alpha(x_{k-1}) = y_k - y_{k-1} = \frac{\alpha(b)-\alpha(a)}{n}.$$

■

Theorem (Rudin 6.9)**.** Let $\alpha : [a,b] \to \mathbb{R}$ be increasing and continuous. Then

(1) If $f : [a,b] \to \mathbb{R}$ is increasing, then $f \in R_\alpha[a,b]$.

(2) If $f : [a,b] \to \mathbb{R}$ is increasing

Proof. Here we will prove (1). The proof of (2) is analogous. First, note that

$$\forall x \in [a,b] \quad f(a) \le f(x) \le f(b) \implies f \text{ is bounded on } [a,b].$$

If $\alpha(a) = \alpha(b)$, then we previously proved $f \in R_\alpha[a,b]$ and $\int_a^b f \, d\alpha = 0$. So, it remains to prove the claim for the case where $\alpha(a) \ne \alpha(b)$. According to the Cauchy Criterion for integrability, in order to show that $f \in R_\alpha[a,b]$, it suffices to show that

$$\forall \varepsilon > 0 \; \exists P \in \Pi \text{ such that } U(f,\alpha,P) - L(f,\alpha,P) < \varepsilon.$$

Let $\varepsilon > 0$ be given. Choose $n \in \mathbb{N}$ be large enough so that $\frac{\alpha(b)-\alpha(a)}{n}[f(b) - f(a)] < \varepsilon$. Let $\tilde{P} = \{x_0, x_1, \dots, x_n\}$ be a partition of $[a,b]$ such that

$$\forall 1 \le k \le n \quad \Delta\alpha_k = \alpha(x_k) - \alpha(x_{k-1}) = \frac{\alpha(b)-\alpha(a)}{n}.$$

We claim that $\tilde{P}$ can be used as the P that we were looking for. Now, since f is increasing, we know that for each $1 \leq k \leq n$

$$M_k = \sup_{x \in [x_{k-1}, x_k]} f(x) = f(x_k)$$

and

$$m_k = \inf_{x \in [x_{k-1}, x_k]} f(x) = f(x_{k-1}).$$

Hence, we see that

$$\begin{aligned} U(f, \alpha, \tilde{P}) - L(f, \alpha, \tilde{P}) &= \sum_{k=1}^{n} (M_k - m_k) \Delta\alpha_k \\ &= \frac{\alpha(b) - \alpha(a)}{n} \sum_{k=1}^{n} [f(x_k) - f(x_{k-1})] \\ &= \frac{\alpha(b) - \alpha(a)}{n} [f(b) - f(a)] < \varepsilon \end{aligned}$$

as desired. ■

> **Theorem** (Rudin 6.10). Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Suppose that f has only finitely many points of discontinuity
> $$y_1 < y_2 < \cdots < y_m$$
> and $\alpha : [a, b] \to \mathbb{R}$ is increasing and α is continuous at $y_1, y_2, \ldots, y_m$. Then $f \in R_\alpha[a, b]$.

Proof. According to the Cauchy Criterion, it suffices to show that

$$\forall \varepsilon > 0 \ \exists P \in \Pi \text{ such that } U(f, \alpha, P) - L(f, \alpha, P) < \varepsilon.$$

Let $\varepsilon > 0$ be given. Let $\tilde{M} = \sup_{x \in [a,b]} |f(x)|$. Let

$$\hat{\varepsilon} = \frac{\varepsilon}{[\alpha(b) - \alpha(a) + 2\tilde{M} + 1]}.$$

We will make the following two claims:

(1) There exists many disjoint intervals $[u_1, v_1], \ldots, [u_m, v_m]$ such that

   (I) $\forall 1 \leq j \leq m$ $y_j \in [u_j, v_j]$.

   (II) $\forall 1 \leq j \leq m$ if $y_j \notin \{a, b\}$, then $y_j \in (u_j, v_j)$

   (III) $\forall 1 \leq j \leq m$ $\alpha(v_j) - \alpha(u_j) < \frac{\hat{\varepsilon}}{m}$ and so

$$\sum_{j=1}^{m} \alpha(v_j) - \alpha(u_j) < \hat{\varepsilon}.$$

(2) Let $K = [a, b] \setminus \bigcup_{j=1}^{m} (u_j, v_j)$. Then f is uniformly continuous on K.

The two claims above will be proven as lemmas after the proof of this theorem. For now, we will assume that the two claims hold.

By claim 2, we know there exists $\delta > 0$ such that for all $s, t \in K$ if $|s - t| < \delta$, then

$$|f(s) - f(t)| < \hat{\varepsilon}.$$

Now, we form a partition $\tilde{P}$ of $[a, b]$ as follows:

(i) $\forall 1 \leq j \leq m$ $u_j, v_j \in \tilde{P}$.

(ii) $\forall 1 \leq j \leq m$ no point of the segment (u_j, v_j) is in $\tilde{P}$

(iii) If $1 \leq k \leq m$ is such that $x_{k-1} \notin \{u_1, \ldots, u_m\}$, then we will choose x_k such that $x_k - x_{k-1} < \delta$.

We claim that this $\tilde{P}$ can be used as the P that we were looking for. Indeed, define the two sets

$$A = \{k : x_{k-1} \notin \{u_1, \ldots, u_m\}\} \quad \text{and} \quad B = \{1, \ldots, n\} \setminus A.$$

For the case that $k \in A$, $x_k - x_{k-1} < \delta$, so for all $s, t \in [x_{k-1}, x_k]$, if $|s - t| < \delta$, then $|f(s) - f(t)| < \hat{\varepsilon}$. Then taking the supremum, we have

$$\sup_{s,t \in [x_{k-1}, x_k]} |f(s) - f(t)| \leq \hat{\varepsilon}$$

and so from lemma 2, we have

$$M_k - m_k \leq \hat{\varepsilon}.$$

If $k \in B$, then

$$M_k - m_k = \sup_{s,t \in [x_{k-1}, x_k]} |f(s) - f(t)| \leq 2\tilde{M}.$$

Therefore,

$$\begin{aligned}
U(f, \alpha, P) - L(f, \alpha, P) &= \sum_{k=1}^{n} (M_k - m_k)\Delta\alpha_k \\
&= \sum_{k \in A} (M_k - m_k)\Delta\alpha_k + \sum_{k \in B} (M_k - m_k)\Delta\alpha_k \\
&\leq \sum_{k \in A} \hat{\varepsilon}\Delta\alpha_k + 2\tilde{M} \sum_{k \in B} \Delta\alpha_k \\
&\leq \hat{\varepsilon}[\alpha(b) - \alpha(a)] + 2\tilde{M}\hat{\varepsilon} \\
&= [\alpha(b) - \alpha(a) + 2\tilde{M}]\hat{\varepsilon} \\
&< \varepsilon.
\end{aligned}$$

■

Lemma. There exists finitely many disjoint intervals

$$[u_1, v_1], \ldots, [u_m, v_m]$$

in $[a, b]$ such that

(1) $\forall 1 \leq j \leq m$ $y_j \in [u_j, v_j]$;

(2) $\forall 1 \leq j \leq m$ if $y_j \notin \{a, b\}$ then $y_j \in (u_j, v_j)$;

(3) $\forall 1 \leq j \leq m$ $\alpha(v_j) - \alpha(u_j) < \frac{\hat{\varepsilon}}{m}$ and so

$$\sum_{j=1}^{m} [\alpha(v_j) - \alpha(u_j)] < \hat{\varepsilon}.$$

Proof. Since for each $1 \leq j \leq m$, α is continuous at y_j, we can choose $\delta_j > 0$ such that

$$\text{if } |y - y_j| < \delta_j, \text{ then } |\alpha(y) - \alpha(y_j)| < \frac{\hat{\varepsilon}}{2m}.$$

Now, let

$$\tilde{\delta} = \frac{1}{4} \min\{\delta_1, \delta_2, \ldots, \delta_m, y_2 - y_1, y_3 - y_2, \ldots, y_m - y_{m-1}\}.$$

For each $1 \leq j \leq m$, we define

(1) If $y_j \notin \{a, b\}$, then $[u_j, v_j] = [y_j - \hat{\delta}, y_j + \hat{\delta}]$

(2) If $y_j = a$, then $[u_j, v_j] = [a, a + \hat{\delta}]$

(3) If $y_j = b$, then $[u_j, v_j] = [b - \hat{\delta}, b]$.

Clearly, there intervals satisfy all the requirements, in particular,

$$\begin{aligned}\alpha(v_j) - \alpha(u_j) &= |\alpha(v_j) - \alpha(u_j)| \\ &\leq |\alpha(v_j) - \alpha(y_j)| + |\alpha(y_j) - \alpha(u_j)| \\ &< \frac{\hat{\varepsilon}}{2m} + \frac{\hat{\varepsilon}}{2m} \\ &= \frac{\hat{\varepsilon}}{m}\end{aligned}$$

where $|v_j - y_j| \leq \hat{\delta} < \delta_j$ and $|u_j - y_j| \leq \hat{\delta} < \delta_j$. ■

Lemma (Claim 2)**.** Let $K = [a,b] \setminus \bigcup_{j=1}^m (u_j, v_j)$. Then f is uniformly continuous on K.

Proof. Note that $\bigcup_{k=1}^m (u_j, v_j)$ is open. Hence,

$$K = [a,b] \setminus \bigcup_{j=1}^m (u_j, v_j) = [a,b] \cap \Big[\bigcup_{j=1}^m (u_j, v_j) \Big]^c$$

is closed. Since $K \subseteq [a,b]$, K is closed, and $[a,b]$ is compact, it follows from the fact that closed subsets of a compact set are compact that K is compact. Since $f : K \to \mathbb{R}$ is continuous and K is compact, we can conclude that f is uniformly continuous on K. ■

Remark (Why is $f : K \to \mathbb{R}$ is continuous?)**.** We will consider four claims:

(1) Suppose f is continuous at a and b. In this case by removing $\bigcup_{j=1}^m (u_j, v_j)$, the discontinuities of f will be removed.

(2) Since f is discontinuous at a, but continuous at b. In this case, by removing $\bigcup_{j=1}^m (u_j, v_j)$ all discontinuities will be removed except a. In this case, removing (u_1, v_1) makes a an isolated point of K. Every function is continuous at every is isolated point of its domain.

(3) Suppose f is continuous at a and discontinuities at b.

(4) Suppose f is both discontinuous at a and b.

Case (3) and (4) follows similarly from case (2).

Theorem (Rudin 6.11)**.** Let $f \in R_\alpha[a,b]$, for all $x \in [a,b]$ $m \leq f(x) \leq M$, $\varphi : [m, M] \to \mathbb{R}$ is continuous. Then $h : \varphi \circ f : [a,b] \to \mathbb{R}$, then $h \in R_\alpha[a,b]$.

Proof. Firs note that a composition of bounded functions is bounded. So $h : \varphi \circ f$ is a bounded function on $[a,b]$. According to the Cauchy criterion, in order to show $h \in R_\alpha[a,b]$, it suffices to show that for all $\varepsilon > 0$, there exists $P \in \Pi$ such that

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

Let $\varepsilon > 0$ be given. Let $\tilde{M} = \sup_{x \in [a,b]} |h(x)|$. Let

$$\hat{\varepsilon} = \frac{\varepsilon}{[\alpha(b) - \alpha(a) + 2\tilde{M} + 1]}.$$

We have

(I) Since φ is continuous in $[m, M]$ and $[m, M]$ is compact, it follows that φ is uniformly continuous on $[, m, M]$. So,

$$\exists 0 < \delta < \hat{\varepsilon} \text{ such that } \forall s, t \in [m, M] \text{ if} |s - t| < \delta \text{ then } |\varphi(s) - \varphi(t)| < \hat{\varepsilon}.$$

(II) Since $f \in R_\alpha[a,b]$, we know from the CauchyCriterion that

$$\exists \tilde{P} \in \Pi \text{ such that } U(f,\alpha,\tilde{P}) - L(f,\alpha,\tilde{P}) = \sum_{k=1}^{n}(M_k - m_k)\Delta\alpha_k < \delta^2.$$

We claim that this $\tilde{P}$ can be used as the P that we were looking for. Indeed, let for all $1 \le k \le n$

$$m_k^* = \inf_{x\in[x_{k-1},x_k]} h(x) \quad\text{and}\quad M_k^* = \sup_{x\in[x_{k-1},x_k]} h(x).$$

Note that

$$U(h,\alpha,\tilde{P}) - L(h,\alpha,\tilde{P}) = \sum_{k=1}^{n}(M_k^* - m_k^*)\Delta\alpha_k.$$

In what follows, we will show that the sum above is less than ε. Divide the indices $1,\ldots,n$ in two classes, namely

$$A = \{k : M_k - m_k < \delta\} \quad\text{and}\quad B = \{k : M_k - m_k \ge \delta\}.$$

We have

$$U(h,\alpha,\tilde{P}) - L(h,\alpha,\tilde{P}) = \sum_{k=1}^{n}(M_k^* - m_k^*)\Delta\alpha_k = \sum_{k\in A}(M_k^* - m_k^*)\Delta\alpha_k + \sum_{k\in B}(M_k^* - m_k^*)\Delta\alpha_k. \tag{1}$$

(*) If $k \in A$, then for all $x, y \in [x_{k-1}, x_k]$, we have

$$\begin{aligned} M_k - m_k < \delta &\implies \sup_{x,y\in[x_{k-1},x_k]} |f(x) - f(y)| \\ &\implies |f(x) - f(y)| < \delta \\ &\implies |\varphi(f(x)) - \varphi(f(y))| < \hat{\varepsilon} \\ &\implies |h(x) - h(y)| < \hat{\varepsilon} \\ &\implies \sup_{x,y\in[x_{k-1},x_k]} |h(x) - h(y)| \le \hat{\varepsilon} \\ &\implies M_k^* - m_k^* \le \hat{\varepsilon}. \end{aligned} \tag{2}$$

(*) For $k \in B$,

$$\begin{aligned} \delta \sum_{k\in B}\Delta\alpha_k = \sum_{k\in B}\delta\Delta\alpha_k &\le \sum_{k\in B}(M_k - m_k)\Delta\alpha_k \\ &\le \sum_{k=1}^{n}(M_k - m_k)\Delta\alpha_k = U(f,\alpha,\tilde{P}) - L(f,\alpha,\tilde{P}) < \delta^2. \end{aligned} \tag{3}$$

It follows from (1), (2), and (3) that

$$\begin{aligned} \sum_{k=1}^{n}(M_k^* - m_k^*)\Delta\alpha_k &= \sum_{k\in A}(M_k^* - m_k^*)\Delta\alpha_k + \sum_{k\in B}(M_k^* - m_k^*)\Delta\alpha_k \\ &\le \sum_{k\in A}\hat{\varepsilon}\Delta\alpha_k + \sum_{k\in B}2\tilde{M}\Delta\alpha_k \\ &= \hat{\varepsilon}\sum_{k=1}^{n}\Delta\alpha_k + 2\tilde{M}\hat{\varepsilon} \\ &= \hat{\varepsilon}[\alpha(b) - \alpha(a)] + 2\tilde{M}\hat{\varepsilon} \\ &= [\alpha(b) - \alpha(a) + 2\tilde{M}]\hat{\varepsilon} \\ &= [\alpha(b) - \alpha(a) + 2\tilde{M}] \cdot \frac{\varepsilon}{\alpha(b) - \alpha(a) + 2\tilde{M} + 1} < \varepsilon \end{aligned}$$

as desired.

■

Chapter 5

Week 5

5.1 Lectures 11-12

5.1.1 Plan

(1) Sequential Criterion for integrability;

(2) Algebraic properties of R.S integral;

(3) Order properties of R.S integrals;

(4) Mean Value Theorem and Generalized Mean Value Theorem for integrals;

(5) Additivity for R.S integrals.

Theorem (Sequential Criterion for R.S Integrability)**.** Let $f : [a, b] \to \mathbb{R}$ be a bounded function and $\alpha : [a, b] \to \mathbb{R}$ is an increasing function. Then

(1) If $f \in R_\alpha[a, b]$, then there exists a sequence of partitions $(P_n)_{n\geq 1}$ in $\Pi[a, b]$ such that $\lim\limits_{n\to\infty}[U(f, \alpha, P_n)- L(f, \alpha, P_n)] = 0$.

(2) If there exists a sequence of partitions $(P_n)_{n\geq 1}$ in $\Pi[a, b]$ such that $\lim\limits_{n\to\infty}[U(f, \alpha, P_n)-L(f, \alpha, P_n)] = 0$, then $f \in R_\alpha[a, b]$, and

$$\int_a^b f\ d\alpha = \lim_{n\to\infty kU(f,\alpha,P_n)} = \lim_{n\to\infty} L(f, \alpha, P_n).$$

Proof. (1) Using the Cauchy Criterion, we see that $f \in R_\alpha[a, b]$ if and only if for all $\varepsilon > 0$, there exists $P_\varepsilon \in \Pi[a, b]$ such that

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

In particular, we can inductively construct a sequence of partitions $(P_n)_{n\geq 1}$ in the following way: for all $n \in \mathbb{N}$, let $\varepsilon = \frac{1}{n}$. Then there exists $P_n \in \Pi$ such that

$$0 \leq U(f, \alpha, P_n) - L(f, \alpha, P_n) < \frac{1}{n}.$$

From the squeeze theorem, it follows that

$$\lim_{n\to\infty}[U(f, \alpha, P_n) - L(f, \alpha, P_n)] = 0.$$

(2) According to Cauchy Criterion, it suffices to show that

$$\forall\varepsilon > 0\ \exists P \in \Pi[a, b] \text{ such that } U(f, \alpha, P) - L(f, \alpha, P) < \varepsilon.$$

Since $\lim_{n\to\infty}[U(f,\alpha,P_n)-L(f,\alpha,P_n)]=0$, there exists an $N\in\mathbb{N}$ such that

$$\forall n>N \quad U(f,\alpha,P_n)-L(f,\alpha,P_n)<\varepsilon.$$

In particular, P_{N+1} can be used as the P that we were looking for. It remains to show that

$$\int_a^b f\ d\alpha=\lim_{n\to\infty}U(f,\alpha,P_n)=\lim_{n\to\infty}L(f,\alpha,P_n).$$

We have for all $n\geq 1$,

$$0\leq U(f,\alpha,P_n)-U(f,\alpha)\leq U(f,\alpha,P_n)-L(f,\alpha)\leq U(f,\alpha,P_n)-L(f,\alpha,P_n).$$

Using the squeeze theorem on the inequality above, we have

$$\lim_{n\to\infty}[L(f,\alpha)-L(f,\alpha,P_n)]=0.$$

So,

$$\lim_{n\to\infty}L(f,\alpha,P_n)=L(f,\alpha)=\int_a^b f\ d\alpha.$$

■

Theorem (Algebraic Properties of R.S Integral)**.** Assume $f,g\in R_\alpha[a,b]$. Then

(i) $\forall k\in\mathbb{R}$, $kf\in R_\alpha[a,b]$ with

$$\int_a^b kf\ d\alpha=k\int_a^b f\ d\alpha$$

;

(ii) $f+g\in R_\alpha[a,b]$ with

$$\int_a^b f+g\ d\alpha=\int_a^b f\ d\alpha+\int_a^b g\ d\alpha.$$

(iii-1) $f^2\in R_\alpha[a,b]$;

(iii-2) $fg\in R_\alpha[a,b]$;

(iv-1) if $g\neq 0$ on $[a,b]$ and $\frac{1}{g}$ is bounded on $[a,b]$, then $\frac{1}{g}\in R_\alpha[a,b]$;

(iv-2) if $g\neq 0$ on $[a,b]$ and $\frac{1}{g}$ is bounded on $[a,b]$, then $\frac{1}{g}\in R_\alpha[a,b]$.

Lemma (lemma 3)**.** Let A be a subset of $\mathbb{R}$ and $f,g:A\in\mathbb{R}$ be two bounded functions. Then

(i) $\sup_A(f+g)\leq\sup_A f+\sup_A g$;

(ii) $\inf_A(f+g)\geq\inf_A f+\inf_A g$;

(iii-1) $\forall k\geq 0$, $\sup_A(kf)=k\sup_A f$;

(iii-2) $\forall k\geq 0\ \inf_A kf=k\inf_A f$;

(iv-1) $\forall k<0\ \sup_A kf=k\inf_A f$;

(iv-2) $\forall k<0\ \inf_A kf=k\sup_A f$;

(v) $\sup_{x,y\in A}|f(x)-f(y)|=\sup_A f-\inf_A f$;

(vi) If there exists a constant $k>0$ such that

$$\forall z,w\in A \quad |f(z)-f(w)|\leq k|g(z)-g(w)|,$$

then

$$\sup_A f - \inf_A f \leq k[\sup_A g - \inf_A g].$$

Lemma (lemma 4). Let $f, g : [a,b] \to \mathbb{R}$ be two bounded functions, $\alpha : [a,b] \to \mathbb{R}$ is an increasing function, and $P \in \Pi[a,b]$. Then

(i) $U(f+g,\alpha,P) \leq U(f,\alpha,P) + U(g,\alpha,P)$;

(ii) $L(f+g,\alpha,P) \geq L(f,\alpha,P) + U(g,\alpha,P)$;

(iii-1) $\forall k \geq 0\ U(kf,\alpha,P) = kU(f,\alpha,P)$

(iii-2) $\forall k \geq 0$, $L(kf,\alpha,P) = kL(f,\alpha,P)$;

(iv-1) $\forall k < 0\ U(f,\alpha,P) = kL(f,\alpha,P)$

(iv-2) $\forall k < 0\ L(kf,\alpha,P) = kU(f,\alpha,P)$.

Theorem (Order Properties of R.S Integral). Assume $f, g \in R_\alpha[a,b]$. Then

(i) If $m \leq f(x) \leq M$ for all $x \in [a,b]$, then

$$m(\alpha(b) - \alpha(a)) \leq \int_a^b f\ d\alpha \leq M(\alpha(b) - \alpha(a)).$$

(ii) If $f \leq g$ on $[a,b]$, then

$$\int_a^b f\ d\alpha \leq \int_a^b g\ d\alpha.$$

Proof. (i) Note that for any $P \in \Pi[a,b]$, we have

$$\int_a^b f\ d\alpha = L(f,\alpha) \geq L(f,\alpha,P)$$
$$\int_a^b f\ d\alpha = U(f,\alpha) \leq U(f,\alpha,P).$$

In particular, for the partition $P = \{a,b\}$, we have

$$\int_a^b f\ d\alpha \geq L(f,\alpha,P) = \Big(\inf_{x\in[a,b]} f(x)\Big)(\alpha(b) - \alpha(a)) \geq m(\alpha(b) - \alpha(a)) \tag{1}$$
$$\int_a^b f\ d\alpha \leq U(f,\alpha,P) = \Big(\sup_{x\in[a,b]} f(x)\Big)(\alpha(b) - \alpha(a)) \leq M(\alpha(b) - \alpha(a)). \tag{2}$$

Using (1) and (2), we obtain our desired result.

(ii) Let $h = g - f$. We have $h \geq 0$, so, by part (i), we have

$$0(\alpha(b) - \alpha(a)) \leq \int_a^b h\ d\alpha.$$

Therefore,

$$0 \leq \int_a^b h\ d\alpha = \int_a^b g - f\ d\alpha = \int_a^b g\ d\alpha - \int_a^b f\ d\alpha$$
$$\implies \int_a^b f\ d\alpha \leq \int_a^b g\ d\alpha.$$

■

Theorem (Triangle Inequality of Integrals). Assume $f \in R_\alpha[a,b]$. Then

(i) $|f| \in R_\alpha[a,b]$;

(ii) $\left|\int_a^b f \, d\alpha\right| \leq \int_a^b |f| \, d\alpha$.

Proof. (i) Define $\varphi : \mathbb{R} \to \mathbb{R}$ by $\varphi(x) = |x|$ which is clearly continuous on $\mathbb{R}$. Since $f \in R_\alpha[a,b]$, it follows from Rudin 6.11 that $\varphi \circ f \in R_\alpha[a,b]$. Hence, we have $|f| \in R_\alpha[a,b]$.

(ii) Recall that

$$|t| \leq s \Longleftrightarrow -s \leq t \leq s.$$

So, our goal is to show that

$$-\int_a^b |f| \, d\alpha \leq \int_a^b f \, d\alpha \leq \int_a^b |f| \, d\alpha.$$

Also, we have

$$-|f(x)| \leq f(x) \leq |f(x)| \quad \forall x \in [a,b].$$

So,

$$-\int_a^b |f(x)| \, d\alpha \leq \int_a^b f(x) \, d\alpha \leq \int_a^b |f(x)| \, d\alpha$$

as desired.

■

Theorem (Mean Value Theorem for Integrals). Let $f : [a,b] \to \mathbb{R}$ be a continuous function, $\alpha : [a,b] \to \mathbb{R}$ is an increasing function and $\alpha(a) \neq \alpha(b)$. Then there exists $c \in [a,b]$ such that

$$f(c) = \frac{1}{\alpha(b) - \alpha(a)} \int_a^b f \, d\alpha.$$

Proof. Since f is continuous on $[a,b]$ and $[a,b]$ is a compact interval in $\mathbb{R}$, it follows from the Extreme Value Theorem that f attains its max and min on $[a,b]$. Let

$$m = \min_{x \in [a,b]} f(x) \quad \text{and} \quad M = \max_{x \in [a,b}]f(x).$$

We have, for all $x \in [a,b]$, $m \leq f(x) \leq M$. Thus,

$$m(\alpha(b) - \alpha(a)) \leq \int_a^b f(x) \, d\alpha \leq M(\alpha(b) - \alpha(a)).$$

Hence,

$$m \leq \frac{1}{\alpha(b) - \alpha(a)} \int_a^b f \, d\alpha \leq M.$$

Using the Intermediate Value Theorem, we see from the assumption that f being continuous on $[a,b]$ that

$$\exists c \in [a,b] \text{ such that } f(c) = \frac{1}{\alpha(b) - \alpha(a)} \int_a^b f \, d\alpha.$$

■

Theorem (Generalized Mean Value Theorme for Integrals). Let $f : [a,b] \to \mathbb{R}$ be continuous, $\alpha : [a,b] \to \mathbb{R}$ is increasing, and $g \in R_\alpha[a,b]$ and either $g \geq 0$ on $[a,b]$ or $g \leq 0$ on $[a,b]$. Then

$$\exists c \in [a,b] \text{ such that } \int_a^b fg \, d\alpha = f(c) \int_a^b g \, d\alpha.$$

Theorem (Additivity for R.S Integrals). Let $f : [a, b] \to \mathbb{R}$ be continuous, $\alpha : [a, b] \to \mathbb{R}$ is increasing, $c \in (a, b)$. Then
$$f \in R_\alpha[a, b] \Longleftrightarrow (f \in R_\alpha[a, c] \text{ and } f \in R_\alpha[c, b]).$$
In this case, we have
$$\int_a^b f \, d\alpha = \int_a^c f \, d\alpha + \int_c^b f \, d\alpha.$$

5.2 Lectures 13-14

5.2.1 Topics

- Theorem: For "nice" α we have $\displaystyle\int_a^b f \, d\alpha = \int_a^b f(x)\alpha'(x) \, dx$;
- Theorem (change of variable)
- The Fundamental Theorem of Calculus
- Integration By Parts
- Unit step function, representing sums by R.S integrals

Lemma. Let $f : [a, b] \to \mathbb{R}$ be a bounded function, $\alpha : [a, b] \to \mathbb{R}$ is an increasing function, $P = \{x_0, x_2, \ldots, x_n\}$ is a partition of $[a, b]$ and $R \in \mathbb{R}$. Then

(1) If for all tags $(s_k)_{1\le k\le n}$ of P, we have $\sum_{k=1}^n f(s_k)\Delta\alpha_k \le R$, then $U(f, \alpha, P) \le R$.

(2) If for all tags $(s_k)_{1\le k\le n}$ of P, we have $R \le \sum_{k=1}^n f(s_k)\Delta\alpha_k$, then $R \le L(f, \alpha, P)$.

Proof. (1) If α is constant, then
$$\sum_{k=1}^n f(s_k)\Delta\alpha_k = 0$$
which implies
$$U(f, \alpha, P) = \sum_{k=1}^n M_k\Delta\alpha_k = 0.$$
So, we may assume that $\alpha(a) \ne \alpha(b)$. It is suffices to show that
$$\forall\varepsilon > 0 \;\; U(f, \alpha, P) \le R + \varepsilon.$$
Let $\varepsilon > 0$ be given. For each $k \in \{1, \ldots, n\}$, we have
$$M_k = \sup_{x\in[x_{k-1},x_k]} f(x) \Longrightarrow \exists s_k \in [x_{k-1}, x_k] \text{ such that } M_k - \frac{\varepsilon}{\alpha(b) - \alpha(a)} < f(s_k).$$
We have
$$\begin{aligned} U(f, \alpha, P) - \sum_{k=1}^n M_k\Delta\alpha_k &< \sum_{k=1}^n \Big[f(s_k) + \frac{\varepsilon}{\alpha(b) - \alpha(a)}\Big]\Delta\alpha_k \\ &= \sum_{k=1}^n f(s_k)\Delta\alpha_k + \frac{\varepsilon}{\alpha(b) - \alpha(a)}\sum_{k=1}^n \Delta\alpha_k \\ &\le R + \varepsilon \end{aligned}$$
as desired.

(2) Completely analogous to (1).

■

Theorem (Rudin 6.17). Let $f : [a,b] \to \mathbb{R}$ be a bounded function, $\alpha : [a,b] \to \mathbb{R}$ be an increasing function, and $\alpha' \in R[a,b]$. Then

$$f \in R_\alpha[a,b] \iff f\alpha' \in R[a,b]$$

and in this case,

$$\int_a^b f \ d\alpha = \int_a^b f(x)\alpha'(x) \ dx.$$

Proof. It suffices to show that

$$\begin{aligned} U(f,\alpha) &= U(f\alpha'), \\ L(f,\alpha) &= L(f\alpha') \end{aligned}$$

Indeed, if we prove (*), then

$$\begin{aligned} f \in R_\alpha[a,b] &\iff U(f,\alpha) = L(f,\alpha) \\ &\iff U(f\alpha') = L(f\alpha') \\ &\iff f\alpha' \in R[a,b]. \end{aligned}$$

Moreover, (*) would imply that

$$\int_a^b f \ d\alpha = U(f,\alpha) = U(f\alpha') = \int_a^b f(x)\alpha'(x) \ dx.$$

In what follows, we will prove $U(f,\alpha) = U(f\alpha')$. The proof of $L(f,\alpha) = L(f\alpha')$ is analogous. Let $P = \{x_0, x_1, \ldots, x_n\}$ be any partition of $[a,b]$. Let (s_k) be any tag of P. Note that by the Mean Value Theorem, we can find a $t_k \in (x_{k-1}, x_k)$ for all $1 \le k \le n$ such that

$$\begin{aligned} \Delta\alpha_k &= \alpha(x_k) - \alpha(x_{k-1}) \\ &= \alpha'(t_k)(x_k - x_{k-1}). \end{aligned}$$

We have

$$\begin{aligned} &\Big|\sum_{k=1}^n f(s_k)\Delta\alpha_k - \sum_{k=1}^n f(s_k)\alpha'(s_k)\Delta x_k\Big| \\ &= \Big|\sum_{k=1}^n f(s_k)\alpha'(t_k)\Delta x_k - \sum_{k=1}^n f(s_k)[\alpha'(t_k) - \alpha'(s_k)]\Delta x_k\Big| \\ &\le \sum_{k=1}^n |f(s_k)||\alpha'(t_k) - \alpha'(s_k)|\Delta x_k \\ &\le \hat{M}\sum_{k=1}^n |\alpha'(t_k) - \alpha'(s_k)|\Delta x_k && (\hat{M} = \sup_{x\in[a,b]}|f(x)|) \\ &\le \hat{M}\sum_{k=1}^n [\sup_{I_k}\alpha' - \inf_{I_k}\alpha']\Delta x_k && \text{(lemma 1)} \\ &= \hat{M}[U(\alpha',P) - L(\alpha',P)]. \end{aligned}$$

Hence, we have

$$\Big|\sum_{k=1}^n f(s_k)\Delta\alpha_k - \sum_{k=1}^n f(s_k)\alpha'(s_k)\Delta x_k\Big| \le \hat{M}[U(\alpha',P) - L(\alpha',P)].$$

Therefore,

$$\begin{aligned}\sum_{k=1}^{n} f(s_k)\Delta\alpha_k &\le \sum_{k=1}^{n} f(s_k)\alpha'(s_k)\Delta x_k + \hat{M}[U(\alpha', P) - L(\alpha', P)] \\ &\le U(f\alpha', P) + \hat{M}[U(\alpha', P) - L(\alpha', P)]. \qquad (1)\end{aligned}$$

By Lemma 5, we have

$$U(f\alpha', P) \le U(f, \alpha, P) + \hat{M}[U(\alpha', P) - L(\alpha', P)]. \qquad (3)$$

It follows from (1) and (2) that

$$|U(f, \alpha, P) - U(f\alpha', P)| \le \hat{M}[U(\alpha', P) - L(\alpha', P)].$$

Note that

$$\begin{aligned}U(f, \alpha) &= \inf_{P\in\Pi} U(f, \alpha, P) \Longrightarrow \exists(P_n^{(1)}) \subseteq \Pi \text{ such that} \\ U(f\alpha') &= \inf_{P\in\Pi} U(f\alpha', P) \Longrightarrow \exists(P_n^{(2)}) \subseteq \Pi \text{ such that } \lim_{n\to\infty} U(f\alpha', P_n^{(2)}) = U(f\alpha'). \qquad (2)\end{aligned}$$

Since $\alpha' \in R[a, b]$, there exists $(P_n^{(3)}) \subseteq \Pi$ such that

$$\lim_{n\to\infty} [U(\alpha', P_n^{(3)}) - L(\alpha', P_n^{(3)})] = 0.$$

Now, for each $n \in \mathbb{N}$, let $P_n = P_n^{(1)} \cup P_n^{(2)} \cup P_n^{(3)}$. We have

$$\begin{aligned}&\forall n \ge 1 \;\; U(f, \alpha) \le U(f, \alpha, P_n) \le U(f, \alpha, P_n^{(1)}) \Longrightarrow \lim_{n\to\infty} U(f, \alpha, P_n) = U(f, \alpha) \qquad (4) \\ &\forall n \ge 1 \;\; U(f\alpha') \le U(f\alpha', P_n) \le U(f\alpha', P_n^2) \Longrightarrow \lim_{n\to\infty} U(f\alpha', P_n) = U(f\alpha'). \qquad (5)\end{aligned}$$

Since P_n is a refinement of $P_n^{(3)}$, we have

$$\begin{aligned}&0 \le [U(\alpha', P_n) - L(\alpha', P_n)] \le U(\alpha', P_n^{(3)} - L(\alpha' P_n^{(3)})) \\ &\Longrightarrow \lim_{n\to\infty} U(\alpha', P_n) - L(\alpha', P_n) = 0. \qquad (6)\end{aligned}$$

It follows from (3) that

$$\forall n \ge 1 \;\; 0 \le U(f, \alpha, P_n) - U(f\alpha', P_n) \le \hat{M}[U(\alpha', P_n) - L(\alpha', P_n)]$$

Applying the squeeze theorem as $n \to \infty$ to both sides of the inequality above, we have

$$\begin{aligned}&\lim_{n\to\infty} |U(f, \alpha, P_n) - U(f\alpha', P_n)| = 0 \\ &\Longrightarrow \left| \lim_{n\to\infty} (U(f, \alpha, P_n) - U(f\alpha', P_n)) \right| = 0 \\ &\Longrightarrow |U(f, \alpha) - U(f\alpha')| = 0 \\ &\Longrightarrow U(f, \alpha) - U(f\alpha') = 0 \\ &\Longrightarrow U(f, \alpha) = U(f\alpha')\end{aligned}$$

■

Theorem (Rudin 6.19; Change of Variable)**.** Let $f \in R_\alpha[a, b]$ and $\varphi : [A, B] \to [a, b]$ be an onto and strictly increasing function. If we let $g = f \circ \varphi$ and $\beta = \alpha \circ \varphi$, then

$$g \in R_\beta[A, B] \;\text{ and }\; \int_a^b f \, d\alpha = \int_A^B g \, d\beta.$$

Proof. Since $\alpha : [a,b] \to \mathbb{R}$ is increasing and $\varphi : [A,B] \to [a,b]$ is an increasing, we see that $\beta = \alpha \circ \varphi$ is also increasing on $[A,B]$. Also, note that, there is a one-to-one correspondence between $\Pi[a,b]$ and $\Pi[A,B]$:

$$H : \Pi[a,b] \to \Pi[A,B]$$

where $P = \{x_0, x_1, \ldots, x_n\}$ corresponding to $[a,b]$ gets mapped to $Q = \{y_0, y_1, \ldots, y_n\}$ corresponding to $[A,B]$. Under this 1-1 correspondence, we have

$$\forall 1 \le k \le n \ \ \varphi([y_{k-1}, y_k]) = [x_{k-1}, x_k].$$

and

$$\forall 1 \le k \le n \ \ \Delta\beta_k = \beta(y_k) - \beta(y_{k-1}) = \alpha(\varphi(y_k)) - \alpha(\varphi(y_{k-1})) = \alpha(x_k) - \alpha(x_{k-1}) = \Delta\alpha_k.$$

and

$$\begin{aligned}\forall 1 \le k \le n \ \ M_k^{(g)} = \sup_{y \in [y_{k-1}, y_k]} g(y) &= \sup_{y \in [y_{k-1}, y_k]} f \circ \varphi(y) \\ &= \sup_{x \in [x_{k-1}, x_k]} f(x) \\ &= M_k^{(f)}\end{aligned}$$

Thus, under the correspondence above, we have

$$U(f, \alpha, P) = U(g, \beta, Q)$$
$$L(f, \alpha, P) = L(g, \beta, Q)$$

where $H(P) = Q$. In order to show $g \in R_\beta[A,B]$, by the Cauchy-Criterion, it suffices to show that

$$\forall \varepsilon > 0 \ \ \exists Q \in \Pi[A,B] \text{ such that } U(g, \beta, Q) - L(g, \beta, Q) < \varepsilon.$$

Let $\varepsilon > 0$ be given. Since $f \in R_\alpha[a,b]$, there exists $P \in \Pi[a,b]$ such that

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

We claim that $Q = H(P)$ can be used as the partition that we were looking for. Indeed,

$$U(g, \beta, H(P)) - L(g, \beta, H(P)) - L(g, \beta, H(P)) = U(f, \alpha, P) - L(f, \alpha, P) < \varepsilon$$

as desired. Also,

$$\begin{aligned}\int_a^b f \, d\alpha = L(f, \alpha) &= \sup_{P \in \Pi[a,b]} L(f, \alpha, P) \\ &= \sup_{Q \in \Pi[a,b]} L(g, \beta, Q) \\ &= L(g, \beta) \\ &= \int_A^B g \, d\beta.\end{aligned}$$

■

Theorem (Fundamental Theorem of Calculus I and II). **(Part I)** Let $f : [a,b] \to \mathbb{R}$ be integrable and $F : [a,b] \to \mathbb{R}$ satisfies $F'(x) = f(x)$ for all $x \in [a,b]$. Then

$$\int_a^b f \, dx = F(b) - F(a).$$

(Part II) Let $g : [a,b] \to \mathbb{R}$ be Riemann Integrable and $G : [a,b] \to \mathbb{R}$ is defined by $G(x) = \int_a^x g(t) \, dt$. Then

(1) G is (uniformly) continuous on $[a,b]$;

(2) If g is continuous at a point $c \in [a,b]$, then G is differentiable at the point c and $G'(c) = g(c)$

(In particular, if $g : [a,b] \to \mathbb{R}$ is continuous, then $G(x) = \int_a^x g(t)\ dt$ is an antiderivative of g on $[a,b]$).

Proof. (I) In what follows, we will show that

$$\forall P \in \Pi[a,b] \quad L(f,P) \le F(b) - f(a) \le U(f,P). \tag{*}$$

Note that as a consequence of (*):

(i) $F(b) - F(a)$ is an upper bound for $\{L(f,P) : P \in \Pi\}$. So,

$$\sup_{P\in\Pi} L(f,P) \le F(b) - F(a) \Longrightarrow L(f) \le F(b) - F(a).$$

(ii) $F(b) - F(a)$ is a lower bound for $\{U(f,P) : P \in \Pi\}$. So,

$$\inf_{P\in\Pi} U(f,P) \ge F(b) - F(a) \Longrightarrow U(f) \ge F(b) - F(a).$$

Thus,

$$L(f) \le F(b) - F(a) \le U(f).$$

Since $f \in R[a,b]$, $L(f) = U(f) = \int_a^b f$, so

$$\int_a^b f \le F(b) - F(a) \le \int_a^b f.$$

Therefore,

$$\int_a^b f = F(b) - F(a)$$

which is our desired result.

Let $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a,b]$. Since $F' = f$ on $[a,b]$, we can use the Mean Value Theorem to find a $t_k \in (x_{k-1}, x_k)$ such that

$$F'(t_k) = f(t_k) = \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}.$$

Hence, we have

$$\begin{aligned} F(b) - F(a) &= \sum_{k=1}^{n} [F(x_k) - F(x_{k-1})] = \sum_{k=1}^{n} F'(t_k)(x_k - x_{k-1}) \\ &= \sum_{k=1}^{n} f(t_k)\Delta x_k. \end{aligned}$$

Therefore, it follows from

$$L(f,P) = \sum_{k=1}^{n} m_k \Delta x_k \le \sum_{k=1}^{n} f(t_k)\Delta x_k \le \sum_{k=1}^{n} M_k \Delta x_k = U(f,P)$$

that

$$L(f,P) \le F(b) - F(a) \le U(f,P).$$

(II) (i) Our goal is to show G is uniformly continuous on $[a,b]$. That is, we need to show

$$\forall \varepsilon > 0\ \exists \delta > 0 \text{ such that } \forall x,y \in [a,b] \text{ if } |x-y| < \delta \text{ then } |G(x) - G(y)| < \varepsilon.$$

Let $\varepsilon > 0$ be given. Let $x, y \in [a,b]$. If $x \ge y$, then

$$\begin{aligned} |G(x) - G(y)| &= \Big| \int_a^x g(t)\ dt - \int_a^y g(t)\ dt \Big| = \Big| \int_y^x g(t)\ dt \Big| \\ &\le \int_y^x |g(t)|\ dt \le \int_y^x R\ dt = R(x-y) = R|x-y| \end{aligned}$$

where $R = \sup_[t \in [a,b]]|g(t)|$. If $y > x$, then

$$\begin{aligned} |G(x) - G(y)| = |G(y) - G(x)| &= \Big| \int_a^y g(t)\ dt - \int_a^x g(t)\ dt \Big| \\ &= \Big| \int_x^y g(t)\ dt \Big| \le \int_x^y |g(t)|\ dt \le \int_x^y R\ dt \\ &= R(y-x) = R|x-y|. \end{aligned}$$

Thus, for all $x, y \in [a,b]$, we have

$$|G(x) - G(y)| \le R|x-y|.$$

Hence, to make sure $|G(x) - G(y)|$ is less than ε, it suffices to make $R|x-y|$ less than ε, that is, it is enough to ensure that $|x-y| < \frac{\varepsilon}{R}$. This argument shows that $\delta = \frac{\varepsilon}{R}$ does the job.

(ii) Now, suppose g is continuous at $c \in [a,b]$. Our goal is to show that $G'(c) = g(c)$. That is, we want to show

$$\lim_{x \to c} \frac{G(x) - G(c)}{x - c} = g(c).$$

That is, our goal is to show that

$$\forall \varepsilon > 0\ \exists \delta > 0 \text{ such that if } 0 < |x-c| < \delta \text{ (with } x \in [a,b]) \text{ then } \Big| \frac{G(x) - G(c)}{x-c} - g(c) \Big| < \varepsilon.$$

Let $\varepsilon > 0$ be given. Since g is continuous at c, for this given ε, there exists $\hat{\delta} > 0$ such that if

$$|t - c| < \hat{\delta} \text{ (with } t \in [a,b]) \text{ then } |g(t) - g(c)| < \frac{\varepsilon}{2}.$$

We claim that this $\hat{\delta}$ can be used as the δ that we were looking for. Indeed, let $\delta = \hat{\delta}$. We will consider the following two cases: **(Case 1)** Suppose $0 < |x-c| < \hat{\delta}$, $x \in [a,b]$, $x > c$. We have

$$\begin{aligned} &\Big| \frac{G(x) - G(c)}{x-c} - g(c) \Big| = \Big| \frac{G(x) - G(c) - g(c)(x-c)}{x-c} \Big| \\ &= \Big| \frac{\int_a^x g(t)\ dt - \int_a^c g(t)\ dt - \int_c^x g(c)\ dt}{x-c} \Big| \\ &= \Big| \frac{1}{x-c} \Big(\int_c^x g(t)\ dt - \int_c^x g(c)\ dt \Big) \Big| \\ &= \Big| \frac{1}{x-c} \int_c^x [g(t) - g(c)]\ dt \Big| \\ &= \frac{1}{|x-c|} \Big| \int_c^x [g(t) - g(c)]\ dt \Big| \\ &\le \frac{1}{x-c} \int_c^x |g(t) - g(c)|\ dt \\ &\le \frac{1}{x-c} \int_c^x \frac{\varepsilon}{2}\ dt \\ &= \frac{1}{x-c} \frac{\varepsilon}{2} (x-c) \\ &= \frac{\varepsilon}{2} \\ &< \varepsilon \end{aligned}$$

as desired. On the other hand, if $0 < |x-c| < \hat{\delta}$, $x \in [a,b]$, and $x < c$, then a similar argument shows our desired result.

■

Theorem (Integration by Parts). Let $u : [a,b] \to \mathbb{R}$ and $v : [a,b] \to \mathbb{R}$ are differentiable and let $u' \in R[a,b]$ and $v' \in R[a,b]$. Then we have

(1) $uv' \in R[a,b]$

(2) $u'v \in R[a,b]$

(3) $\displaystyle\int_a^b uv' \, dx = u(b)v(b) - u(a)v(a) - \int_a^b u'v \, dx.$

Proof. (1) Since $u : [a,b] \to \mathbb{R}$ is differentiable, we have $u \in C[a,b]$. So, we have $u \in R[a,b]$. By assumption, $v' \in R[a,b]$ and so we can conclude that $uv' \in R[a,b]$.

(2) Using the same argument above, we have $uv' \in R[a,b]$.

(3) By the product rule, we have
$$(uv)' = u'v + uv'.$$
In particular, since $(uv)'$ is a sum of integrable functions, it belongs to $R[a,b]$. Now, we integrate both sides
$$\int_a^b (uv)' \, dx = \int_a^b u'v \, dx + \int_a^b uv' \, dx. \tag{I}$$
According to FTC I, we have
$$\int_a^b (uv)' \, dx = [uv]_{x=a}^{x=b} = u(b)v(b) - u(a)v(a). \tag{II}$$
Hence, we have (I) and (II) imply that
$$u(b)v(b) - u(a)v(a) = \int_a^b u'v \, dx + \int_a^b uv' \, dx$$
which further implies that
$$\int_a^b uv' \, dx = u(b)v(b) - u(a)v(a) - \int_a^b u'v \, dx.$$
■

Definition (Unit Step Function). The **unit step function** $I : \mathbb{R} \to \mathbb{R}$ is defined by
$$I(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if} x > 0 \end{cases}.$$

Remark. Note that for all $s \in |R$, we have
$$I(x-s) = \begin{cases} 0 & \text{if } x \leq s \\ 1 & \text{if} x > s \end{cases}.$$
Also, for all $c \neq 0$, we have
$$cI(x-s) = \begin{cases} 0 & \text{if } x \leq s \\ c & \text{if } x > s. \end{cases}$$

Theorem (Rudin 6.15). Let $f : [a,b] \to \mathbb{R}$ be a bounded function, $s \in (a,b)$, f is right continuous at s, and $\alpha(x) = I(x-s)$. Then

$$f \in R_\alpha[a, b] \text{ and } \int_a^b f \, d\alpha = f(s).$$

Proof. see hw4 ■

Theorem (Rudin 6.16). Suppose for all $n \geq 1$, $c_n \geq 0$, $\sum_{n=1}^{N} c_n$ converges, $s_1 < s_2 < s_3 < \cdots$ are points in (a, b), $\alpha : [a, b] \to \mathbb{R}$ is continuous. Then

$$f \in R_\alpha[a, b] \text{ and } \int_a^b f \, d\alpha = \sum_{n=1}^{\infty} c_n f(s_n).$$

Proof. See hw4. ■

Chapter 6

Week 6

6.1 Plan

- Uniform convergence
- Uniform convergence and boundedness
- Uniform convergence and continuity
- Cauchy criterion for uniform convergence
- Uniform convergence and differentiability
- Uniform Convergence and integrability

Definition (Uniform Convergence)**.** We say $(f_n : A \to \mathbb{R})$ **converges uniformly** to $f : A \to \mathbb{R}$ if

$$\forall \varepsilon > 0 \ \exists N_\varepsilon \text{ such that } \forall n > N_\varepsilon \ \forall x \in A \ |f_n(x) - f(c)| < \varepsilon.$$

Theorem (Uniform Convergence Preserves Boundedness)**.** Let $A \neq \emptyset$, for each $n \in \mathbb{N}$, $f_n : A \to \mathbb{R}$ is bounded, and $f_n \to f$ uniformly on A. Then $f : A \to \mathbb{R}$ is bounded.

Remark. Please make a clear distinction between the following statements:

(1) For all $n \in \mathbb{N}$, $f_n : A \to \mathbb{R}$ is bounded:

$$\forall n \in \mathbb{N} \ \exists \hat{M}_n \text{ such that } |f_n(c)| \leq \hat{M}_n.$$

(2) For all $(f_n)_{n \geq 1}$ is uniformly bounded:

$$\exists M \text{ such that } \forall n \geq 1 \ \forall x \in A \ \ |f_n(x)| \leq M.$$

(3) $(f_n)_{n \geq 1}$ is pointwise bounded:

$$\forall x \in A \ (f_n(x))_{n \geq 1} \text{ is bounded}.$$

Proof. Our goal is to show that

$$\exists M \in \mathbb{R} \text{ such that } \forall x \in A \ |f(x)| \leq M.$$

Since $f_n \to f$ uniformly on A, we have for all $\varepsilon > 0$, there exists N such that for all $n > N$ and for all $x \in A$,

$$\begin{aligned}
&|f_n(x) - f(x)| < \varepsilon \\
&\implies |f(x)| - |f_n(x)| < \varepsilon \\
&\implies |f(x)| < \varepsilon + |f_n(x)|.
\end{aligned}$$

In particular, for $\varepsilon = 1$, there exists $N \in \mathbb{N}$ such that

$$\forall n > N \ \forall x \in A \ |f(x)| < 1 + |f_n(x)|.$$

Now, if we let $n = N + 1$, we get

$$\forall x \in A \quad |f(x)| < 1 + |f_{N+1}(x)|. \tag{1}$$

Since, by assumption, f_{N+1} is bounded, there exists a number $\hat{M}_{N+1}$ such that

$$\forall x \in A \quad |f_{N+1}(x)| \leq \hat{M}_{N+1}. \tag{2}$$

It follows from (1) and (2) that

$$\forall x \in A \quad |f(x)| < 1 + \hat{M}_{N+1}.$$

Clearly, we can use $1 + \hat{M}_{N+1}$ as the same M we were looking for. ■

Theorem (Rudin 7.12)**.** Let $A \subseteq (X, d)$ and $x \in A$. Suppose for all $n \in \mathbb{N}$, $f_n : A \to \mathbb{R}$ is continuous at c and $f_n \to f$ uniformly on A. Then $f : A \to \mathbb{R}$ is continuous at c.

Proof. Our goal is to show that

$$\forall \varepsilon > 0 \ \exists \delta > 0 \text{ if } d(x, c) < \delta \text{ then } |f(x) - f(c)| < \varepsilon.$$

Let $\varepsilon > 0$ be given. Since $f_n \to f$ uniformly on A, there exists $N \in \mathbb{N}$ such that for all $n > N$ and for all $z \in A$, we have

$$|f_n(z) - f(z)| < \frac{\varepsilon}{3}. \tag{1}$$

Also, since f_{N+1} is continuous at c,

$$\exists \hat{\delta} > 0 \text{ such that } \forall x \in N_{\hat{\delta}}(c) \cap A \quad |f_{N+1}(x) - f_{N+1}(c)| < \frac{\varepsilon}{3}. \tag{2}$$

We claim that $\hat{\delta} > 0$ can be used as the same $\hat{\delta}$ that we were looking for. Indeed, for all $x \in N_{\hat{\delta}}(c) \cap A$, we have

$$\begin{aligned}
|f(x) - f(c)| &\leq |f(x) - f_{N+1}(x)| + |f_{N+1}(x) - f_{N+1}(c)|f_{N+1}(c) - f(c)| \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\
&= \varepsilon
\end{aligned}$$

as desired. ■

Remark (A Useful Observation)**.** Let (a_n) be a sequence of real numbers. Suppose $a_n \to a$ in $\mathbb{R}$. Suppose there exists N such that

$$\forall m, n > N \quad |a_n - a_m| < \frac{1}{3}.$$

So, by taking the limit as $n \to \infty$, it follows from the order limit theorem that for each $n > N$, we have

$$\lim_{m \to \infty} |a_n - a_m| \leq \lim_{m \to \infty} \frac{1}{3} = \frac{1}{3}.$$

More generally, given $\varepsilon > 0$, if there exists N such that

$$\forall n, m > N \quad |a_n - a_m| \leq \varepsilon.$$

We will use the remark above to prov the following theorem:

Theorem (Cauchy Criterion for Uniform Convergence)**.** Let $A \neq \emptyset$ and suppose for each $n \in \mathbb{N}$, $f_n : A \to \mathbb{R}$ is a sequence of functions. Then $(f_n)_{n\geq 1}$ converges uniformly on A if and only if for all $\varepsilon > 0$, there exists N such that for all $m, n > N$ and for all $x \in A$, $|f_n(x) - f_m(x)| < \varepsilon$.

Proof. ($\Longrightarrow$) Suppose there exists $f : A \to \mathbb{R}$ such that $f_n \to f$ uniformly on A. Our goal is to find an N such that for all $m, n > N$ and for all $x \in A$

$$|f_n(x) - f_m(x)| < \varepsilon. \tag{*}$$

Since $f_n \to f$ uniformly on A, for the given $\varepsilon > 0$, there exists $\hat{N}$ such that

$$\forall k > \hat{N} \quad \forall x \in A \quad |f_k(x) - f(x)| < \frac{\varepsilon}{2}.$$

do We claim that this $\hat{N}$ can be used as the N that we were looking for. Indeed, if $m, n > \hat{N}$ and $x \in A$, then

$$\begin{aligned} |f_n(x) - f_m(x)| &\leq |f_n(x) - f(x)| + |f_m(x) - f(x)| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon \end{aligned}$$

as desired.

(($\Longleftarrow$)) Suppose for all $\varepsilon > 0$, there exists N such that for all $n, m > N$ and for all $x \in A$

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

Our goal is to show that $(f_n)_{n\geq 1}$ converges uniformly on A. It follows from the assumption that at each point $x \in A$, the sequence of real numbers $(f_n(x))_{n\geq 1}$ is a Cauchy sequence in $\mathbb{R}$. Since $\mathbb{R}$ is complete, we can conclude that at each point $x \in A$, the sequence of real numbers $(f_n(x))_{n\geq 1}$ that converges. This tells us that the sequence of functions $(f_n)_{n\geq 1}$ is pointwise convergent on A. Let's denote the pointwise limit of $(f_n)_{n\geq 1}$ by $f : A \to \mathbb{R}$. In what follows, we will prove that $f_n \to f$ uniformly on A. To this end, we need to show

$$\forall \varepsilon > 0 \ \exists N \text{ such that } \forall n > N \ \forall x \in A \ |f_n(x) - f(x)| < \varepsilon.$$

Let $\varepsilon > 0$ be given. It follows from the assumption that for the given $\varepsilon > 0$, there exists $\hat{N}$ such that

$$\forall m, n > \hat{N} \ \forall x \in A \ |f_n(x) - f_m(x)| < \frac{\varepsilon}{2}.$$

We claim that this $\hat{N}$ can be used as the N we were looking for. Indeed, if $n > \hat{N}$ and $x \in A$, then

$$\forall m > \hat{N} \quad |f_n(x) - f_m(x)| < \frac{\varepsilon}{2}.$$

So, by taking the limit as $m \to \infty$ (using the Useful Observation), we have

$$|f_n(x) - f(x)| \leq \frac{\varepsilon}{2} < \varepsilon$$

as desired. ∎

Theorem (Rudin 7.17)**.** Suppose for each $n \in \mathbb{N}$, $f_n : [a, b] \to \mathbb{R}$ is a sequence of differentiable functions and $f_n \to f$ pointwise. Assume that f_n' converges uniformly to a function g on $[a, b]$. Then f is differentiable to a function g on $[a, b]$.

Proof. Our goal is to show that for all $c \in [a,b]$, $f'(c) = g(c)$. Let $c \in [a,b]$. We want to show that

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

That is, we want to show that for all $\varepsilon > 0$, there exists $\delta > 0$ such that if $0 < |x-c| < \delta$ (with $x \in [a,b]$), then

$$\Big|\frac{f(x) - f(c)}{x - c} - g(c)\Big| < \varepsilon.$$

To this end, let $\varepsilon > 0$ be given. Since $f_n' \to g$ uniformly, we can find an N_1 such that for all $n > N_1$, for all $z \in [a,b]$, $|f_n' - g(z)| < \frac{\varepsilon}{3}$. This tells us that (f_n') fulfills the Cauchy Criterion for Uniform Convergence and so there exists an N_2 such that for all $m, n > N_2$ and for all $z \in [a,b]$, we have

$$|f_n'(z) - f_m'(z)| < \frac{\varepsilon}{3}.$$

Let $N = \max\{N_2, N_2\} + 1$. Also, f_N is differentiable at c, so for our given ε, there exists $\hat{\delta} > 0$ such that if $0 < |x - c| < \hat{\delta}$ (with $x \in [a,b]$)

$$\Big|\frac{f_N(x) - f_N(c)}{x - c} - f_N'(c)\Big| < \frac{\varepsilon}{3}.$$

We claim that this $\hat{\delta}$ can be used as the δ that we were looking for. Indeed, if $x \in [a,b]$ and $0 < |x-c| < \hat{\delta}$, then

$$\begin{aligned}
\Big|\frac{f(x) - f(c)}{x - c} - g(c)\Big| &= \Big|\frac{f(x) - f(c)}{x - c} - \frac{f_N(x) - f_N(c)}{x - c} + \frac{f_N(x) - f_N(c)}{x - c} - f_N'(c) + f_N'(c) - g(c)\Big| \\
&\le \Big|\frac{f(x) - f(c)}{x - c} - \frac{f_N(x) - f_N(c)}{x - c}\Big| + \Big|\frac{f_N(x) - f_N(c)}{x - c} - f_N'(c)\Big| + |f_N'(c) - g(c)| \\
&< \Big|\frac{(f - f_N)(x) - (f - f_N)(c)}{x - c}\Big| + \frac{\varepsilon}{3} + \frac{\varepsilon}{3}.
\end{aligned}$$

To complete the proof, it suffices to show that the first term on the second inequality above is less than $\frac{\varepsilon}{3}$. Suppose without loss of generality that $x < c$ where $x \in [a,b]$. Then for every $m > N$, we can apply the Mean Value Theorem to the function $f_m - f_N$ on the interval $[x, c]$. That is, for all $m > N$, there exists $\alpha_m \in (x, c)$ such that

$$(f_m - f_N)'(\alpha_m) = \frac{(f_m - f_N)(c) - (f_m - f_N)(x)}{c - x}.$$

By (2), we know that $|f_m'(\alpha_m) - f_N'(\alpha_m)| < \frac{\varepsilon}{3}$. So, we have

$$\Big|\frac{(f_m - f_N)(c) - (f_m - f_N)(x)}{c - x}\Big| < \frac{\varepsilon}{3}.$$

By taking the limit as $m \to \infty$, we get

$$\Big|\frac{(f - f_N)(c) - (f - f_N)(x)}{c - x}\Big| \le \frac{\varepsilon}{3}.$$

So,

$$\Big|\frac{(f - f_N)(x) - (f - f_N)(c)}{x - c}\Big| \le \frac{\varepsilon}{3}$$

as desired. ■

Lemma (lemma 1)**.** Let A be nonempty. Let $f : A \to \mathbb{R}$. Suppose $(f_n : A \to \mathbb{R})$ is a sequence of functions. The following statements are equivalent:

(1) $f_n \to f$ uniformly on A;

(2) $\lim\limits_{n \to \infty} \sup_{x \in A} |f_n(x) - f(x)| = 0$

Lemma (lemma 2)**.** Let $\alpha : [a,b] \to \mathbb{R}$ is increasing, $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ are bounded, and $f \le g$. Then

$$L(f,\alpha) \le L(g,\alpha) \text{ and } U(f,\alpha) \le U(g,\alpha).$$

Theorem (Rudin 7.16)**.** Let $\alpha : [a,b] \to \mathbb{R}$ is increasing, for each $n \in \mathbb{N}$ $f_n \in R_\alpha[a,b]$, and $f_n \to f$ uniformly on $[a,b]$. Then $f \in R_\alpha[a,b]$ and

$$\int_a^b f \; d\alpha = \lim_{n\to\infty} \int_a^b f_n \; d\alpha.$$

Proof. Since uniform convergence preserves boundedness, we can conclude that $f : [a,b] \to \mathbb{R}$ is bounded. Now, in order to show that $f \in R_\alpha[a,b]$, it suffices to show that $L(f,\alpha) = U(f,\alpha)$. For each $n \in \mathbb{N}$, let

$$r_n = \sup_{x\in[a,b]} |f_n(x) - f(x)|.$$

Since $f_n \to f$ uniformly, we know that $\lim_{n\to\infty} r_n = 0$. For each $n \in \mathbb{N}$, we have

$$r_n = \sup_{x\in[a,b]} |f_n(x) - f(x)| \Longrightarrow |f(x) - f_n(x)| \le r_n \forall x \in [a,b].$$

Hence,

$$\forall x \in [a,b] \quad f_n(x) - r_n \le f(x) \le f_n(x) + r_n. \tag{*}$$

So, it follows from lemma 2 that

$$L(f_n - r_n, \alpha) \le L(f,\alpha) \le U(f,\alpha) \le U(f_n + r_n).$$

Thus,

$$0 \le U(f,\alpha) - L(f,\alpha) \le U(f_n + r_n, \alpha) - L(f_n - r_n, \alpha).$$

Note that

$$\begin{aligned} U(f_n + r_n, \alpha) - L(f_n - r_n, \alpha) &= \int_a^b (f_n + r_n) \; d\alpha - \int_a^b (f_n - r_n) \; d\alpha \\ &= \int_a^b \Big[(f_n + r_n) - (f_n - r_n)\Big] \; d\alpha \\ &= \int_a^b 2r_n \; d\alpha \\ &= 2r_n[\alpha(b) - \alpha(a)]. \end{aligned}$$

So,

$$0 \le U(f,\alpha) - L(f,\alpha) \le 2r_n[\alpha(b) - \alpha(a)].$$

Using the Squeeze Theorem, we have $U(f,\alpha) = L(f,\alpha)$ (by applying the limit as $n \to \infty$). Now, it follows from (*) that

$$\int_a^b (f_n - r_n) \; d\alpha \le \int_a^b f \; d\alpha \le \int_a^b (f_n + r_n) \; d\alpha.$$

So,

$$\int_a^b (-r_n) \; d\alpha \le \int_a^b f \; d\alpha - \int_a^b f_n \; d\alpha \le \int_a^b r_n \; d\alpha.$$

Thus,

$$-r_n[\alpha(b) - \alpha(a)] \le \int_a^b f \; d\alpha - \int_a^b f_n \; d\alpha \le r_n[\alpha(b) - \alpha(a)].$$

Using the Squeeze Theorem as $n \to \infty$, we have

$$\lim_{n\to\infty} \left[\int_a^b f \ d\alpha - \int_a^b f_n \ d\alpha \right] = 0.$$

That is,

$$\lim_{n\to\infty} \int_a^b f_n \ d\alpha = \int_a^b f \ d\alpha.$$

■

Chapter 7

Week 7

7.1 Lectures 17-18

7.1.1 Plan

- Series of functions
- Cauchy Criterion for Uniform Convergence of Series
- Weiertstrass M-Test

Theorem (Term-by-Term Continuity Theorem). Let $A \subseteq (X, d)$ be nonempty. Suppose for all $n \in \mathbb{N}$ $f_n : A \to \mathbb{R}$ is a sequence of continuous functions, and $\sum_{n=1}^{\infty} f_n$ converges uniformly to $f : A \to \mathbb{R}$. Then $f : A \to \mathbb{R}$ is continuous.

Proof. Applying the corresponding theorem for sequences of functions to the sequence of partial sums $s_m = f_1 + \cdots + f_n$. That is,

$$\sum_{n=1}^{\infty} f_n = f \Longrightarrow s_m \to f \text{ uniformly} \Longrightarrow f \text{ is continuous}$$

since s_m is continuous. ■

Theorem (Term-by-Term Differentiability Theorem). Assume for each $n \in \mathbb{N}$, $f_n : [a, b] \to \mathbb{R}$ is a sequence of differentiable functions, $\sum_{n=1}^{\infty} f_n = f$ pointwise on $[a, b]$, and $\sum_{n=1}^{\infty} f_n'$ converges uniformly on $[a, b]$. Then f is differentiable on $[a, b]$ and

$$\left(\sum_{n=1}^{\infty} f_n\right)' = \sum_{n=1}^{\infty} f_n'.$$

Proof. Apply the corresponding theorem for sequences of functions to the sequence of partial sums $s_m = f_1 + \cdots + f_m$. ■

Theorem (Term-by-Term Integrability). Let $\alpha : [a, b] \to \mathbb{R}$ is an increasing function, for each $n \geq 1$, $f_n \in R_\alpha[a, b]$, and $\sum_{n=1}^{\infty} f_n$ converges uniformly to $f : [a, b] \to \mathbb{R}$. Then

$$f \in R_\alpha[a, b] \quad \text{and} \quad \int_a^b \sum_{n=1}^{\infty} f_n \, d\alpha = \sum_{n=1}^{\infty} \int_a^b f_n \, d\alpha.$$

Proof. Apply the corresponding theorem for sequences of functions to the sequence of partial sums $s_m = f_1 + \cdots + f_m$. ■

Theorem (Cauchy Criterion for Uniform Convergence of Series of Functions)**.** Let A be a nonempty set and suppose for each $k \in \mathbb{N}$, $f_k : A \to \mathbb{R}$. Then

$$\sum_{k=1}^{\infty} f_k \text{ converges uniformly if and only if for all } \varepsilon > 0, \text{ there exists an } N \text{ such that for all}$$

$$n > m > N \text{ and for all } x \in A, \left|\sum_{k=1}^{n} f_k(x)\right| < \varepsilon.$$

Theorem (Weierstrass M-Test)**.** Let A be a nonempty set, for all $n \in \mathbb{N}$ $f_n : A \to \mathbb{R}$, for all $n \in \mathbb{N}$, there exists M_n such that for all $x \in A$, $|f_n(x)| \leq M_n$, and $\sum_{n=1}^{\infty} M_n$ converges. Then

$$\sum_{n=1}^{\infty} f_n \text{ converges uniformly on } A.$$

Proof. According to the Cauchy Criterion for uniform convergence of series of functions, it suffices to show that for all $\varepsilon > 0$, there exists N such that for all $n > m > N$ and for all $x \in A$

$$\left|\sum_{k=m+1}^{n} f_k(x)\right| < \varepsilon. \tag{*}$$

Let $\varepsilon > 0$. Note, by assumption, $\sum_{n=1}^{\infty} M_n$ converges. Thus, for our given ε, there exists $\hat{N}$ such that

$$\forall m > m > \hat{N} \quad \left|\sum_{k=m+1}^{n} M_k\right| < \varepsilon.$$

We claim that we can use this $\hat{N}$ as the N that we were looking for. Indeed, if we let $N = \hat{N}$, then (*) will hold because for all $n > m > \hat{N}$ and for all $x \in A$

$$\left|\sum_{k=m+1}^{n} f_k(x)\right| \leq \sum_{k=m+1}^{n} |f_k(x)| \leq \sum_{k=m+1}^{n} M_k < \varepsilon$$

as desired. ■

7.2 Lectures 20-21

7.2.1 Plan

(1) Dini's Theorem

Theorem (Rudin 7.13)**.** Let (X, d) be a metric space, let $K \subseteq X$ be a compact set, and suppose for each $n \in \mathbb{N}$, $f_n : K \to \mathbb{R}$ is continuous. Assume further that $f_n \to f$ pointwise on K where $f : K \to \mathbb{R}$ is continuous, and that for all $n \in \mathbb{N}$, $f_{n+1} \leq f_n$. Then $f_n \to f$ uniformly on K.

Proof. Let $\varepsilon > 0$ be given. Our goal is to show, there exists N such that for all $n > N$ and for all $\in K$, we have

$$|f_n(x) - f(x) < \varepsilon.|$$

For each $n \in \mathbb{N}$, let $g_n = f_n - f$. So, our goal is to show that there exists an N such that

$$\forall n > N \ \forall x \in K \ |g_n(x)| < \varepsilon.$$

First, we observe that for all $g_n \geq 0$. Indeed, we see that for each $x \in K$, $(f_n(x))_{n\geq 1}$ is a decreasing sequence of real numbers that converges to $f(x)$. It follows from the Monotone Convergence Theorem that $f(x) = \inf_{n\in\mathbb{N}} f_n(x)$. Thus, for all $n \in \mathbb{N}$, we have

$$f(x) \leq f_n(x).$$

Therefore, for all $n \in \mathbb{N}$, $g_n \geq 0$. To get our desired result, all we need to show is that there exists an N such that for all $n > N$ and for all $x \in K$, $g_n(x) < \varepsilon$. We can reframe our desired conclusion in the following way:

$$\exists N \in \mathbb{N} \text{ such that } \forall n > N \quad \{x \in K : g_n(x) \geq \varepsilon\} = \emptyset. \tag{*}$$

Let $K_n = g_n^{-1}([\varepsilon, \infty])$ for each $n \in \mathbb{N}$. Our goal is to show that for all $n > N$, $K_n = \emptyset$. Observe further that for each $n \in \mathbb{N}$, K_n is a compact set. Indeed, we see that for each $n \in \mathbb{N}$, $g_n : K \to \mathbb{R}$ is continuous and $[\varepsilon, \infty)$ is a closed set in $\mathbb{R}$. From this, we can see that $K_n = g_n^{-1}([\varepsilon, \infty))$ is closed in K because preimages of closed sets under a continuous map is closed. Thus, we can see that each K_n must be compact because K is compact, $K_n \subseteq K$ and K_n is closed.

For our third observation, we see that $K_{n+1} \subseteq K_n$ for all $n \in \mathbb{N}$. Indeed, we see that for every $x \in K_{n+1}$,

$$g_{n+1}(x) \geq \varepsilon \underbrace{\Longrightarrow}_{g_{n+1}\leq g_n} g_n(x) \geq \varepsilon \Longrightarrow x \in K_n.$$

This tells us that to show (*), it is enough to find an $N \in \mathbb{N}$ such that $K_N = \emptyset$. Assume for contradiction that for all $n \in \mathbb{N}$, $K_n \neq \emptyset$. Because

- $K_{n+1} \subseteq K_n$ for all $n \in \mathbb{N}$;
- $\forall n$, K_n is compact;
- $\forall n \in \mathbb{N} K_n \neq \emptyset$;

we can see, by the Nested Compact Interval Property that

$$\bigcap_{n=1}^{\infty} K_n \neq \emptyset.$$

Therefore, there exists $x \in \bigcap_{n=1}^{\infty} K_n$, that is, there exists an $x \in K$ such that

$$\forall n \in \mathbb{N} \quad g_n(x) \geq \varepsilon.$$

This contradicts the fact that $g_n(x) \to 0$ (Indeed, we can see that this is the case because $f_n \to f$ is pointwise and so $g_n = f_n - f \to 0$ pointwise). ■

Theorem (The Arzela-Ascoli Theorem)**.** Let (X, d) be a metric space, $K \subseteq X$ where K is infinite and compact, and $(f_n : K \to \mathbb{R})_{n\geq 1}$ is uniformly bounded, and $(f_n : K \to \mathbb{R})_{n\geq 1}$ is equicontinuous. Then (f_n) has a uniformly convergent subsequence.

Before proving this remarkable theorem, we would like to go over some key terms defined within the statement above so that we may understand the context better. Below, we note the key differences between continuity and uniformly continuity. First, consider the definitions between the two terms.

Definition (Continuity)**.**

(i) We say that $f : A \to \mathbb{R}$ where $A \subseteq (X, d)$ is continuous if for all $c \in A$, for all $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $x \in A$ if $d(x, c) < \delta$, then

$$|f(x) - f(c)| < \varepsilon.$$

(ii) We say that $f : A \to \mathbb{R}$ is uniformly continuous if for all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x, y \in A$ if $d(x, y) < \delta$, then

$$|f(x) - f(y)| < \varepsilon.$$

Definition (Equicontinuous Sequence of Functions). Let $A \subseteq (X, d)$. A sequence of functions $(f_n : A \to \mathbb{R})_{n\geq 1}$ is said to be **equicontinuous** if

$$\forall \varepsilon > 0 \ \exists \delta > 0 \text{ such that } \forall n \in \mathbb{N} \ \forall x, c \in A \text{ if } d(x, c) < \delta \text{ ,then } |f_n(x) - f_n(c)| < \varepsilon.$$

We will now outline the key main steps taken to prove the Arzela-Ascoli theorem.

(1) We let E be countable dense subset of K. We will use the assumption that (f_n) is uniformly bounded to show that there exists a subsequence (f_{n_k}) of (f_n) that converges at each point of E. To simplify the notation, we let $g_k = f_{n_k}$. The main proof technique for this step is to use Cantor's diagonal process.

(2) We will use the assumption that (f_n) is equicontinuous to prove that the sequence (g_k), which we constructed in step 1 above is uniformly convergent on the entire K. For this step, the idea is to prove that (g_k) satisfies the Cauchy Criterion for Uniform Convergence; that is, we will need to show that

$$\forall \varepsilon > 0 \ \exists N \text{ such that } \forall m, n > N \ \forall x \in K \ |g_m(x) - g_n(x)| < \varepsilon.$$

Note that for each $x \in K$ and $r \in E$, we have

$$|g_m(x) - g_n(x)| \leq |g_m(x) - g_m(r)| + |g_m(r) - g_n(r)| + |g_n(r) - g_n(x)|.$$

The first term and third term on the right-hand side of the inequality above can be made small by using the equicontinuity of (g_k). The middle term can be made small using the assumption that (g_k) converges at $r \in E$ and so $(g_k(r))$ is a Cauchy sequence of real numbers.

Proof of Step 2

Proof. Suppose $(g_k)_{k\geq 1}$ is equicontinuous, for each $r \in E$, the sequence of numbers $(g_k(r))_{k\geq 1}$ converges, and E is a countable dense subset of K. Then $(g_k)_{k\geq 1}$ converges uniformly on K.

To this end, let $\varepsilon > 0$ be given. Our goal is to show to find an N such that

$$\forall m, n > N \ \forall x \in K \ |g_m(x) - g_n(x)| < \varepsilon. \tag{*}$$

Note that

(i) $(g_k)_{k\geq 1}$ is equicontinuous, so for the given $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\forall k \in \mathbb{N} \ \forall x, c \in K \text{ if } d(x, c) < \delta \ |g_k(x) - g_k(c)| < \frac{\varepsilon}{3}.$$

(ii) For each $r \in E$, the sequence of numbers $(g_k(r))_{k\geq 1}$ is convergent so it is Cauchy. Hence,

$$\forall c \in E \ \ \exists N_r \text{ such that } \forall m, n > N_r \ |g_m(r) - g_n(r)| < \frac{\varepsilon}{3}.$$

Notice that (since E is dense in K), we have

$$K \subseteq \bigcup_{r\in E} B_\delta(r)$$

where $B_\delta(r)$ is an open ball of radius δ centered at r. So, $\{B_\delta(r)\}_{r\in E}$ is an open cover of K. Since K is compact, this open cover has a finite subcover, that is, there exists $r_1, \ldots, r_\ell \in E$ such that

$$K \subseteq [B_\delta(r_1) \cup \cdots \cup B_\delta(r_\ell)];$$

that is, every point in K is within δ of at least one of $r_1, \ldots, r_\ell$. We claim that $\max\{N_{r_1}, \ldots, N_{r_\ell}\}$ can be used as the N we were looking for. Indeed, if we let $N = \max\{N_{r_1}, \ldots, N_{r_\ell}\}$, then (*) will hold. The reason is as follows:

Suppose $m, n > N$ and $x \in K$. Since $x \in K$, there exists $i \in \{1, \ldots, \ell\}$ such that $x \in B_\delta(r_i)$. We have

$$\begin{aligned}
|g_m(x) - g_n(x)| &= |g_m(x) - g_m(r_i) + g_m(r_i) - g_n(r_i) + g_n(r_i) - g_n(x)| \\
&\leq |g_m(x) - g_m(r_i)| + |g_m(r_i) - g_n(r_i)| + |g_n(r_i) - g_n(x)| \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon
\end{aligned}$$

as desired.

■

Proof of Step 1

Proof. In what follows, we will prove a more general statement:

Let $E = \{x_i : i \in \mathbb{N}\}$ is some countable subset of K and $(f_n : K \to \mathbb{R})_{n\geq 1}$ is pointwise bounded. Then (f_n) has a subsequence that converges at each point $r \in E$.

- The sequence of real numbers $(f_n(x_1))_{n\geq 1}$ is bounded. Using the Bolzano-Weierstrass theorem, it has a subsequence $(f_{n_k}(x_1))_{k\geq 1}$ that converges. To emphasize that this subsequence is generated by considering the values of x_1, we will use the notation $f_{1,k} = f_{n_k}$.
- The sequence of real numbers $(f_{1,k}(x_2))_{k\geq 1}$ is bounded. Hence, the Bolzano-Weierstrass theorem implies that it has a convergent subsequence $f_{2,k}(x_2)_{k\geq 1}$.
- The sequence of real number $(f_{2,k}(x_3))_{k\geq 1}$. In general, for every $m > 2$, the sequence $(f_{m-1,k})_{k\geq 1}$ has a subsequence $(f_{m,k})_{k\geq 1}$ such that $(f_{m,k}(x_m)_{k\geq 1})$ converges. In this way, we will obtain the following table of functions:

$$\begin{array}{llll}
f_{1,1} & f_{1,2} & f_{1,3} & f_{1,4}\cdots \\
f_{2,1} & f_{2,2} & f_{2,3} & f_{2,4}\cdots \\
f_{3,1} & f_{3,2} & f_{3,3} & f_{3,4}\cdots \\
\vdots & & & \\
f_{m,1} & f_{m,2} & f_{m,3} & f_{m,4}\cdots \\
\vdots & & &
\end{array}$$

Now, consider the diagonal sequence $f_{1,1}, f_{2,2}, f_{3,3}, \ldots$. Let $g_k = f_{k,k}$ for each $k \in \mathbb{N}$. We claim that $(g_k)_{k\geq 1}$ converges at x_i for all $i \in \mathbb{N}$. Indeed, let $i \in \mathbb{N}$. Note that

$$g_{i+1}, g_{i+2}, g_{i+3}, \cdots$$

is a subsequence of $f_{i,1}, f_{i,2}, f_{i,3}, \ldots$. By construction,

$$f_{i,1}(x_i), f_{i,2}(x_i), f_{i,3}(x_i), \ldots$$

converges. Since $g_{i+1}(x_i), g_{i+2}(x_i), g_{i+3}(x_i), \ldots$ is a subsequence of the sequence above, it also converges. Thus,

$$g_1(x_i), g_2(x_i), \ldots, g_i(x_i), g_{i+1}(x_i), g_{i+2}(x_i), g_{i+3}(x_i), \ldots$$

also converges. ■

Chapter 8

Week 8

Chapter 9

Week 9

Chapter 10

Week 10

10.1 Topics

- Additional remarks on the Arzela-Ascoli Theorem.
- $(BC(E; R), d_\infty)$ is a complete metric space.
- Weierstrass Approximation Theorem

10.1.1 Additional Remarks on Arzela-Ascoli

During our proof we only used the fact that the sequence (f_n) is **pointwise bounded** (not the strong assumption of uniform boundedness). So, the Arzela-Ascoli Theorem can be restated as follows:

Theorem (Arzela-Ascoli)**.** Let (X, d) be a metric space, $K \subseteq X$ is an infinite and compact set, $(f_n : K \to \mathbb{R})_{n\geq 1}$ is pointwise bounded, and $(f_n : K \to \mathbb{R})_{n\geq 1}$ is equicontinuous. Then (f_n) has a uniformly convergent subsequence.

Furthermore, the assumption that K is infinite is not needed and that the claim of the theorem still holds even if the compact set K is finite. In this case, we can simply set $E = K$ and adjust the steps accordingly.

Now, let $A \subseteq (X, d)$ be a nonempty set. A family $\mathcal{F}$ of real-valued functions defined on A is said to be **equicontinuous** on A if

$$\forall \varepsilon > 0 \ \exists \delta_\varepsilon \text{ such that } \forall f \in \mathcal{F} \ \forall x, c \in A \text{ if } d(x, c) < \delta_\varepsilon \text{ then } |f(x) - f(c)| < \varepsilon.$$

Let X be any infinite set of functions and equip X with a metric d. Each such metric d introduces a **mode of convergence** for sequences of functions in X. That is, the sequence $(f_n)_{n\geq 1}$ in X converges to $f \in X$ if and only if

$$\forall \varepsilon > 0 \ \exists N \text{ such that } \forall n > N \ d(f_n, f) < \varepsilon.$$

10.1.2 The Space of Bounded and Continuous functions is a Complete

We make the following observations:

- If E is compact, then
$$BC(E; \mathbb{R}) = C(E; \mathbb{R})$$
by the Extreme Value Theorem.
- Suppose (f_n) is a sequence in $BC(E; \mathbb{R})$. We see that
$$f_n \to f \text{ in } BC(E; \mathbb{R}) \text{ if and only if } f_n \to f \text{ uniformly on } E.$$

From this, we can further restate the Arzela-Ascoli as follows:

Theorem (Restating the Arzela-Ascoli Theorem). Let (K, d) be a compact metric space, $X = C(K; \mathbb{R})$ equipped with the d_∞ metric, $(f_n)_{n\geq 1}$ is a sequence in X, and $(f_n)_{n\geq 1}$ is equicontinuous and pointwise bounded. Then $(f_n)_{n\geq 1}$ has a convergent subsequence in X.

Now, we can restate a stronger version of the theorem using all of our observations.

Theorem (A Stronger Version of the Arzela-Ascoli Theorem). Let (K, d) be a compact metric space, $X = C(K; \mathbb{R})$ equipped with the d_∞ metric, and $\mathcal{F} \subseteq X$ is a collection of functions. Then the following two statements are equivalent:

(1) Any sequence $(f_n)_{n\geq 1}$ in $\mathcal{F}$ has a convergent subsquence.

(2) The collection $\mathcal{F}$ is equicontinuous and pointwise bounded.

Theorem (Rudin 7.15). Let (E, d) be a metric space. Let $X = BC(E; \mathbb{R})$ equipped with d_∞ metric. Then (X, d_∞) is a complete metric space.

Proof. Let $(f_n)_{n\geq 1}$ be a Cauchy sequence in $(BC(E; \mathbb{R}), d_\infty)$. Our goal is to show that $(f_n)_{n\geq 1}$ is convergent in $(BC(E; \mathbb{R}), d_\infty)$. Since $(f_n)_{n\geq 1}$ is Cauchy, we see that for all $\varepsilon > 0$, there exists an N such that $\forall n, m > N$ such that

$$d_\infty(f_n, f_m) < \varepsilon \iff \sup_{x\in E} |f_n(x) - f_m(x)| < \varepsilon.$$

Since $|f_n(x) - f_m(x)| \leq \sup_{x\in E} |f_n(x) - f_m(x)|$ for all $x \in E$, we can see that for all $n, m > N$,

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

Hence, we can see that (f_n) fulfills the Cauchy Criterion for uniform convergence; that is, (f_n) converges uniformly on E.

Now, let $f = \lim_{n\to\infty} f_n$. For each n, f_n being bounded and continuous tells us that the limiting function f must also be bounded and continuous since uniform convergence preserves both continuity and boundedness. Hence, we can see that f must belong to $BC(E; \mathbb{R})$ and so $(BC(E; \mathbb{R}))$ must be a complete metric space. ■

The following theorem tells us that that the space of polynomials of degree at most n over any compact interval in $\mathbb{R}$ is dense in the space of continuous functions. That is, any continuous function can be approximated by a sequence of polynomials in $\mathbb{R}$.

10.1.3 Weierstrass Approximation Theorem

Theorem (Rudin 7.26). If $f : [a, b] \to \mathbb{R}$ is continuous, then there exists a sequence $(p_n)_{n\geq 1}$ of polynomials such that $d_\infty(p_n, f) \to 0$.

Before embarking on the proof of this essential theorem, we will give mention to a few facts that will be useful in completing this task.

Definition (Modulus of Continuity). Let $g : [a, b] \to \mathbb{R}$ be a bounded function. The **modulus of continuity** of g is defined by the following equation:

$$\forall 0 < r \leq b - a \quad W_g(r) = \sup_{x,z\in[a,b]} |g(x) - g(z)|$$

with $|x - z| \leq r$.

Proposition (Fact 1)**.** If $g : [a, b] \to \mathbb{R}$ is continuous, then $w(r) \to 0$ as $r \to 0^+$.

Proof. Our goal is to show that for all $\varepsilon > 0$, there exists $\delta > 0$ such that if $0 < r < \delta$, then

$$|w(r) - 0| < \varepsilon;$$

that is, we want to show that

$$\forall \varepsilon > 0\ \exists \delta > 0 \text{ such that if } 0 < r < \delta \text{ then } \sup_{x,z \in [a,b]} |g(x) - g(z)| < \varepsilon.$$

Let $\varepsilon > 0$. To this end, it suffices to show that for all $\varepsilon > 0$, there exists $\delta > 0$ such that if $0 < r < \delta$, for all $x, z \in [a, b]$ with $|x - z| \leq r$, we have

$$|g(x) - g(z)| < \varepsilon.$$

But this is just the direct consequence of uniform continuity of g on $[a, b]$. That is, for all $\varepsilon > 0$ there exists $\delta > 0$ such that for all $x, z \in [a, b]$ if $|x - z| < \delta$, then $|g(x) - g(z)| < \varepsilon$. ■

Proposition (Fact 2)**.** Let $g : [a, b] \to \mathbb{R}$ be a bounded function. Then for all $x, z \in [a, b]$ and for all $0 < r \leq b - a$, we have

$$|g(x) - g(z)| \leq \left[\frac{|x - z|}{r} + 1\right] w(r).$$

Proof. We will consider two cases. Namely, we will consider the case where $|x - z| \leq r$ and $|x - z| > r$ (assume in this case that $x < z$ without loss of generality). If $|x - z| \leq r$, we have

$$|g(x) - g(z)| \leq w(r) \leq \left[\frac{|x - z|}{r} + 1\right] w(r).$$

Now, suppose $|x - z| > r$. Let $m \in \mathbb{N}$ be such that $x + mr \leq z < x + (m + 1)r$. We have

$$\begin{aligned} |g(x) - g(z)| &\leq |g(x) - g(x + r)| + |g(x + r) - g(x + 2r)| + \cdots + |g(x + mr) - g(z)| \\ &\leq w(r) + w(r) + \cdots + w(r) \\ &= (m + 1)w(r) \\ &\leq \left(\frac{|x - z|}{r} + 1\right) w(r). \end{aligned}$$

Note that the last inequality holds because $x + mr \leq z$ implies that $m \leq \frac{z-x}{r}$. ■

Proof of Weierstrass Approximation Theorem

Proof. We will consider two cases; namely, $[a, b] = [0, 1]$ or $[a, b] \neq [0, 1]$. Starting with the first case, our goal is to construct as sequence of polynomials p_n in $\P_n$ for which the result holds.

(1) Consider $n + 1$ equally spread meshpoints in the interval $[0, 1]$ as follows:

$$0 = x_{0,n} < x_{1,n} < \cdots < x_{n,n} = 1$$

with $x_{j,n} = \frac{j}{n}$.

(2) Associate with the jth meshpoint, a polynomial $\varphi_{j,n}(x) \in \P_n$ such that

(I) $\varphi_{j,n}(x) \geq 0$ for all $x \in [0, 1]$

(II) $\displaystyle\sum_{j=0}^{n} \varphi_{j,n}(x) = 1$ for all $x \in [0, 1]$

(III) $\displaystyle\sum_{j=0}^{n} [x - x_{j,n}]^2 \varphi_{j,n}(x) \leq \frac{M}{n}$ for some M independent of x and n.

■