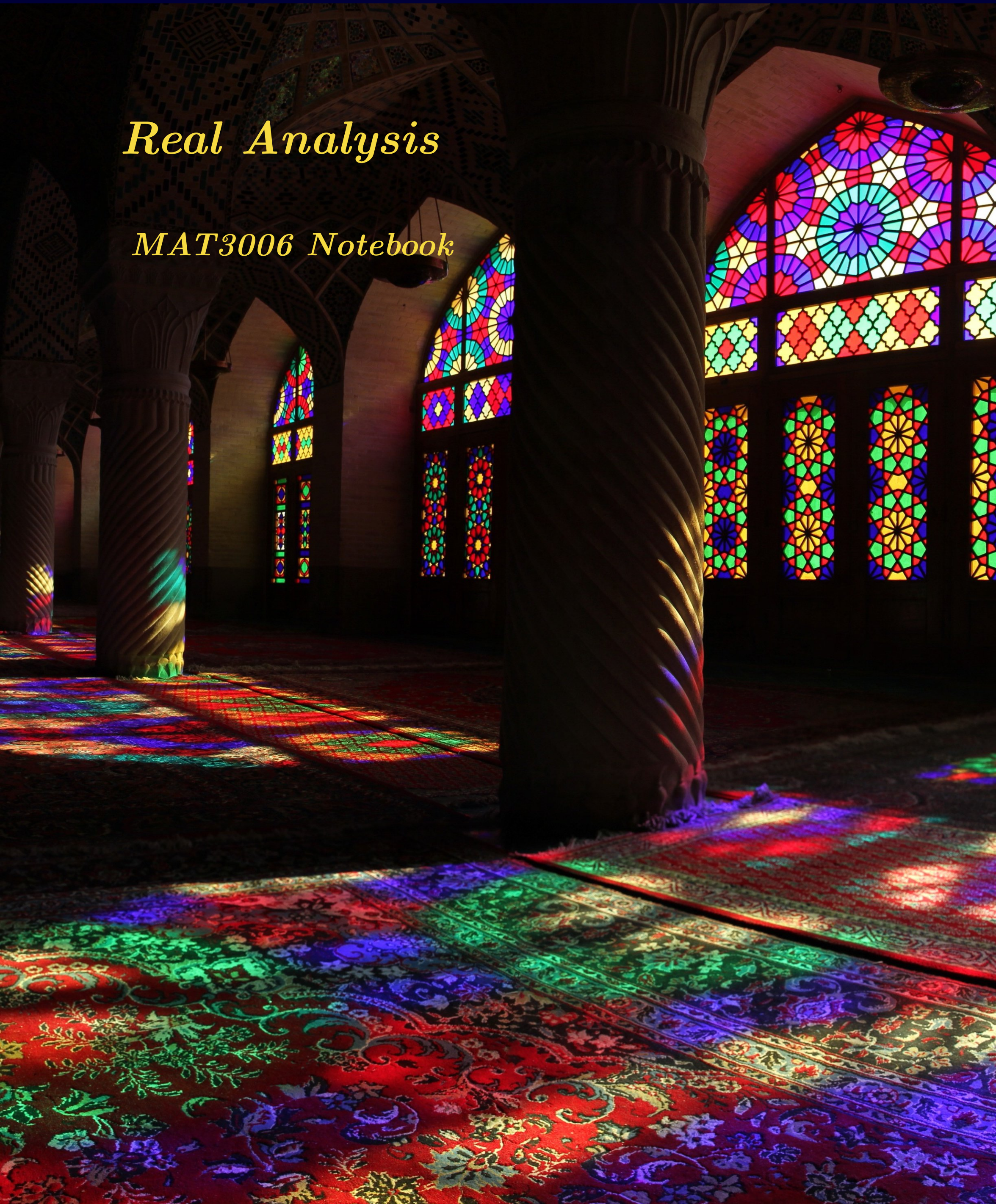
香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

Real Analysis

MAT3006 Notebook



The First Edition

A FIRST COURSE

IN

REAL ANALYSIS

A FIRST COURSE

IN

REAL ANALYSIS

MAT3006 Notebook

Lecturer

Prof. Daniel Wong

The Chinese University of Hongkong, Shenzhen

Tex Written By

Mr. Jie Wang

The Chinese University of Hongkong, Shenzhen

香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

Contents

Acknowledgments

This book is taken notes from the MAT3006 in spring semester, 2019. These lecture notes were taken and compiled in LaTeX by Jie Wang, an undergraduate student in spring 2019. The tex writter would like to thank Prof. Daniel Wong and some students for their detailed and valuable comments and suggestions, which significantly improved the quality of this notebook. Students taking this course may use the notes as part of their reading and reference materials. This version of the lecture notes were revised and extended for many times, but may still contain many mistakes and typos, including English grammatical and spelling errors, in the notes. It would be greatly appreciated if those students, who will use the notes as their reading or reference material, tell any mistakes and typos to Jie Wang for improving this notebook.

Notations and Conventions

$\overline{\mathbb{R}}$	Extended real line
$f_n \rightrightarrows f$	The sequence f_n converges to f uniformly
$\\|\boldsymbol{x}\\|_p$	p-norm of the vector $\boldsymbol{x}$
$C(X)$	Space of continuous functions
E'	The set of accumulation points of E
$\bar{E}$	Closure of E
$\mathcal{P}[a,b]$	The collection of polynomials $p(x)$ with $x \in [a,b]$
$\limsup$	Upper limit
$\liminf$	Lower limit
μ	Measure
$m^*(E)$	Outer measure of E
$\mathcal{M}$	The collection of Lebesgue measurable subsets of $\mathbb{R}$
$m(E)$	Lebesgue measure of E
$\mathbb{P}(E)$	Power set of E
$(\Omega, \mathcal{T}, \mu)$	Measurable space
$\mathcal{X}_E$	Characteristic function on E
$\int_E f(x)\,\mathrm{d}x$	Lebesgue integral of f over the measurable set E
$\int_a^b f(x)\,\mathrm{d}x$	Riemann integral of f over the interval (a,b)
$\mathcal{B}$	Borel σ-algebra
$\mathcal{A} \otimes \mathcal{B}$	The smallest σ-algebra containing $\mathcal{A} \times \mathcal{B}$

1.2. Monday for MAT3006

1.2.1. Overview on uniform convergence

Definition 1.3 [Convergence] Let $f_n(x)$ be a sequence of functions on an interval $I = [a,b]$. Then $f_n(x)$ converges **pointwise** to $f(x)$ (i.e., $f_n(x_0) \to f(x_0)$) for $\forall x_0 \in I$, if

$$\forall \varepsilon > 0, \exists N_{x_0,\varepsilon} \text{ such that } |f_n(x_0) - f(x_0)| < \varepsilon, \forall n \geq N_{x_0,\varepsilon}$$

We say $f_n(x)$ converges uniformly to $f(x)$, (i.e., $f_n(x) \rightrightarrows f(x)$) for $\forall x_0 \in I$, if

$$\forall \varepsilon > 0, \exists N_{\varepsilon} \text{ such that } |f_n(x_0) - f(x_0)| < \varepsilon, \forall n \geq N_{\varepsilon}$$

■ **Example 1.6** It is clear that the function $f_n(x) = \frac{n}{1+nx}$ converges pointwise into $f(x) = \frac{1}{x}$ on $[0,\infty)$, and it is uniformly convergent on $[1,\infty)$.

Proposition 1.2 If $\{f_n\}$ is a sequence of continuous functions on I, and $f_n(x) \rightrightarrows f(x)$, then the following results hold:

1. $f(x)$ is continuous on I.
2. f is (Riemann) integrable with $\int_a^b f_n(x)\,\mathrm{d}x \to \int_a^b f(x)\,\mathrm{d}x$.
3. Suppose furthermore that $f_n(x)$ is **continuously differentiable**, and $f_n'(x) \rightrightarrows g(x)$, then $f(x)$ is differentiable, with $f_n'(x) \to f'(x)$.

We can put the discussions above into the content of series, i.e., $f_n(x) = \sum_{k=1}^n S_k(x)$.

Proposition 1.3 If $S_k(x)$ is continuous for $\forall k$, and $\sum_{k=1}^n S_k \rightrightarrows \sum_{k=1}^\infty S_k$, then

1. $\sum_{k=1}^\infty S_k(x)$ is continuous,
2. The series $\sum_{k=1}^\infty S_k$ is (Riemann) integrable, with $\sum_{k=1}^\infty \int_a^b S_k(x)\,\mathrm{d}x = \int_a^b \sum_{k=1}^\infty S_k(x)\,\mathrm{d}x$
3. If $\sum_{k=1}^n S_k$ is continuously differentiable, and the derivative of which is uniform

convergent, then the series $\sum_{k=1}^{\infty} S_k$ is differentiable, with

$$\left(\sum_{k=1}^{\infty} S_k(x)\right)' = \sum_{k=1}^{\infty} S_k'(x)$$

Then we can discuss the properties for a special kind of series, say power series.

Proposition 1.4 Suppose the power series $f(x) = \sum_{k=1}^{\infty} a_k x^k$ has radius of convergence R, then

1. $\sum_{k=1}^{n} a_k x^k \rightrightarrows f(x)$ for any $[-L, L]$ with $L < R$.
2. The function $f(x)$ is continuous on $(-R, R)$, and moreover, is differentiable and (Riemann) integrable on $[-L, L]$ with $L < R$:

$$\int_0^x f(t)\,\mathrm{d}t = \sum_{k=1}^{\infty} \frac{a_k}{k+1} x^{k+1}$$
$$f'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}$$

1.2.2. Introduction to MAT3006

What are we going to do.

1. (a) Generalize our study of (sequence, series, functions) on $\mathbb{R}^n$ into a metric space.
   (b) We will study spaces outside $\mathbb{R}^n$.

   Remark:

   - For (a), different metric may yield different kind of convergence of sequences. For (b), one important example we will study is $X = C[a,b]$ (all continuous functions defined on $[a,b]$.) We will generalize X into $C_b(E)$, which means the set of bounded continuous functions defined on $E \subseteq \mathbb{R}^n$.
   - The insights of analysis is to find a **unified** theory to study sequences/series on a metric space X, e.g., $X = \mathbb{R}^n, C[a,b]$. In particular, for $C[a,b]$, we will see that
     - most functions in $C[a,b]$ are nowhere differentiable. (repeat part of

content in MAT2006)

- We will prove the existence and uniqueness of ODEs.
- the set poly$[a,b]$ (the set of polynomials on $[a,b]$) is dense in $C[a,b]$. (analogy: $\mathbb{Q} \subseteq \mathbb{R}$ is dense)

2. Introduction to the Lebesgue Integration.

   For convergence of integration $\int_a^b f_n(x)\,\mathrm{d}x \to \int_a^b f(x)$, we need the pre-conditions (a) $f_n(x)$ is continuous, and (b) $f_n(x) \rightrightarrows f(x)$. The natural question is that can we relax these conditions to

   (a) $f_n(x)$ is integrable?

   (b) $f_n(x) \to f(x)$ pointwisely?

   The answer is yes, by using the tool of Lebesgue integration. If $f_n(x) \to f(x)$ and $f_n(x)$ is Lebesgue integrable, then $\int_a^b f_n(x)\,\mathrm{d}x \to \int_a^b f(x)\,\mathrm{d}x$, which is so called the **dominated convergence**.

1.2.3. Metric Spaces

We will study the **length** of an element, or the **distance** between two elements in an arbitrary set X. First let's discuss the length defined on a well-structured set, say vector space.

Definition 1.4 [Normed Space] Let X be a vector space. A **norm** on X is a function $\|\cdot\| : X \to \mathbb{R}$ such that

1. $\|\boldsymbol{x}\| \ge 0$ for $\forall \boldsymbol{x} \in X$, with equality iff $\boldsymbol{x} = \boldsymbol{0}$
2. $\|\alpha\boldsymbol{x}\| = |\alpha|\|\boldsymbol{x}\|$, for $\forall \alpha \in \mathbb{R}$ and $\boldsymbol{x} \in X$.
3. $\|\boldsymbol{x}+\boldsymbol{y}\| \le \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ (triangular inequality)

Any vector space equipped with $\|\cdot\|$ is called a **normed space**.

■ **Example 1.7** 1. For $X = \mathbb{R}^n$, define

$$\|\boldsymbol{x}\|_2 = \left(\sum_{i=1}^n x_i^2\right)^{1/2} \quad \text{(Euclidean Norm)}$$

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} \quad (p\text{-norm})$$

2. For $X = C[a,b]$, define

$$\|f\|_\infty = \max_{x\in[a,b]} |f(x)|$$

$$\|f\|_p = \left(\int_a^b |f(x)|^p \, \mathrm{d}x\right)^{1/p}$$

Exercise: check the norm defined above are well-defined. ■

Here we can define the distance in an arbitrary set:

Definition 1.5 A set X is a **metric space** with metric (X,d) if there exists a (distance) function $d : X \times X \to \mathbb{R}$ such that

1. $d(\boldsymbol{x},\boldsymbol{y}) \ge 0$ for $\forall \boldsymbol{x},\boldsymbol{y} \in X$, with equality iff $\boldsymbol{x} = \boldsymbol{y}$.
2. $d(\boldsymbol{x},\boldsymbol{y}) = d(\boldsymbol{y},\boldsymbol{x})$.
3. $d(\boldsymbol{x},\boldsymbol{z}) \le d(\boldsymbol{x},\boldsymbol{y}) + d(\boldsymbol{y},\boldsymbol{z})$.

■

■ **Example 1.8** 1. If X is a normed space, then define $d(\boldsymbol{x},\boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|$, which is so called the metric induced from the norm $\|\cdot\|$.

2. Let X be any (non-empty) set with $\boldsymbol{x},\boldsymbol{y} \in X$, the discrete metric is given by:

$$d(\boldsymbol{x},\boldsymbol{y}) = \begin{cases} 0, & \text{if } x = y \\ 1, & \text{if } x \ne y \end{cases}$$

Exercise: check the metric space defined above are well-defined. ■

R We will mostly study metric spaces whose metrics come from the norm of a normed space. Adopting the infinite norm discussed in Example (1.7), we can define a metric on $C[a,b]$ by

$$d_\infty(f,g) = \|f-g\|_\infty := \max_{x\in[a,b]} |f(x)-g(x)|$$

which is the correct metric to study the uniform convergence for $\{f_n\} \subseteq C[a,b]$.

Definition 1.6 Let (X,d) be a metric space. An **open ball** centered at $\boldsymbol{x} \in X$ of radius r is the set

$$B_r(\boldsymbol{x}) = \{\boldsymbol{y} \in X \mid d(\boldsymbol{x},\boldsymbol{y}) < r\}.$$

■ **Example 1.9** 1. For $X = \mathbb{R}^2$, we can draw the $B_1(\mathbf{0})$ with respect to the metrics d_1, d_2:

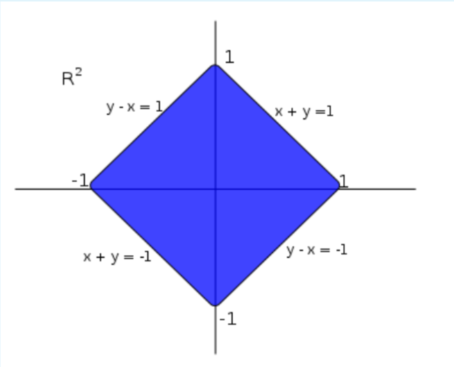


Figure 1.1: $B_1(\mathbf{0})$ w.r.t. the metric d_1

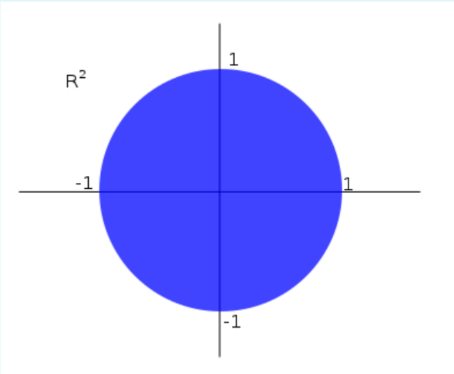


Figure 1.2: $B_1(\mathbf{0})$ w.r.t. the metric d_2

1.5. Wednesday for MAT3006

Reviewing.

- Normed Space: a norm on a vector space
- Metric Space
- Open Ball

1.5.1. Convergence of Sequences

Since $\mathbb{R}^n$ and $C[a,b]$ are both metric spaces, we can study the convergence in $\mathbb{R}^n$ and the functions defined on $[a,b]$ at the same time.

Definition 1.14 [Convergence] Let (X,d) be a metric space. A sequence $\{x_n\}$ in X is **convergent** to x if $\forall \varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$d(x_n,x) < \varepsilon, \forall n \geq N.$$

We can denote the convergence by

$$x_n \to x, \quad \text{or} \quad \lim_{n\to\infty} x_n = x, \quad \text{or} \quad \lim_{n\to\infty} d(x_n,x) = 0$$

Proposition 1.10 If the limit of $\{x_n\}$ exists, then it is unique.

R Note that the proposition above does not necessarily hold for topology spaces.

Proof. Suppose $x_n \to x$ and $x_n \to y$, which implies

$$0 \leq d(x,y) \leq d(x,x_n) + d(x_n,y), \forall n$$

Taking the limit $n \to \infty$ both sides, we imply $d(x,y) = 0$, i.e., $x = y$. ■

■ **Example 1.16** 1. Consider the metric space $(\mathbb{R}^k, d_\infty)$ and study the convergence

$$\begin{aligned}\lim_{n\to\infty} \boldsymbol{x}_n = \boldsymbol{x} &\Longleftrightarrow \lim_{n\to\infty}\left(\max_{i=1\dots,k}|x_{n_i} - x_i|\right) = 0\\ &\Longleftrightarrow \lim_{n\to\infty}|x_{n_i} - x_i| = 0, \forall i = 1,\dots,k\\ &\Longleftrightarrow \lim_{n\to\infty} x_{n_i} = x_i,\end{aligned}$$

i.e., the convergence defined in d_∞ is the same as the convergence defined in d_2.

2. Consider the convergence in the metric space $(C[a,b], d_\infty)$:

$$\begin{aligned}\lim_{n\to\infty} f_n = f &\Longleftrightarrow \lim_{n\to\infty}\left(\max_{[a,b]}|f_n(x) - f(x)|\right) = 0\\ &\Longleftrightarrow \forall\varepsilon > 0, \forall x \in [a,b], \exists N_\varepsilon \text{ such that } |f_n(x) - f(x)| < \varepsilon, \forall n \geq N_\varepsilon\end{aligned}$$

which is equivalent to the uniform convergence of functions, i.e., the convergence defined in d_2.

■

Definition 1.15 [Equivalent metrics] Let d and ρ be metrics on X.

1. We say ρ is **stronger** than d (or d is **weaker** than ρ) if

$$\exists K > 0 \text{ such that } d(x,y) \leq K\rho(x,y), \forall x,y \in X$$

2. The metrics d and ρ are equivalent if there exists $K_1, K_2 > 0$ such that

$$d(x,y) \leq K_1\rho(x,y) \leq K_2 d(x,y)$$

■

(R) The strongerness of ρ than d is depicted in the graph below:

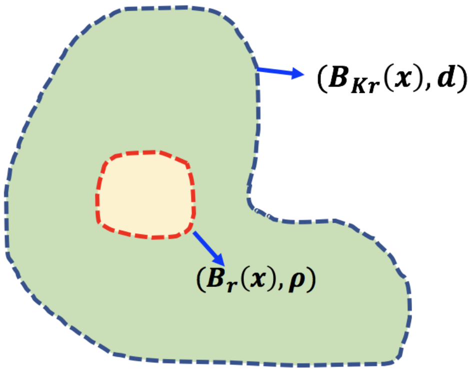


Figure 1.4: The open ball $(B_r(x), \rho)$ is contained by the open ball $(B_{Kr}(x), d)$

For each $x \in X$, consider the open ball $(B_r(x), \rho)$ and the open ball $(B_{Kr}(x), d)$:

$$B_r(x) = \{y \mid \rho(x, y) < r\}, \quad B_{Kr}(x) = \{z \mid d(x, z) < Kr\}.$$

For $y \in (B_r(x), \rho)$, we have $d(x, y) < K\rho(x, y) < Kr$, which implies $y \in (B_{Kr}(x), d)$, i.e, $(B_r(x), \rho) \subseteq (B_{Kr}(x), d)$ for any $x \in X$ and $r > 0$.

Example 1.17 1. d_1, d_2, d_∞ in $\mathbb{R}^n$ are equivalent

$$d_1(\boldsymbol{x}, \boldsymbol{y}) \le d_\infty(\boldsymbol{x}, \boldsymbol{y}) \le n d_1(\boldsymbol{x}, \boldsymbol{y})$$

$$d_2(\boldsymbol{x}, \boldsymbol{y}) \le d_\infty(\boldsymbol{x}, \boldsymbol{y}) \le \sqrt{n} d_2(\boldsymbol{x}, \boldsymbol{y})$$

We use two relation depiected in the figure below to explain these two inequalities:

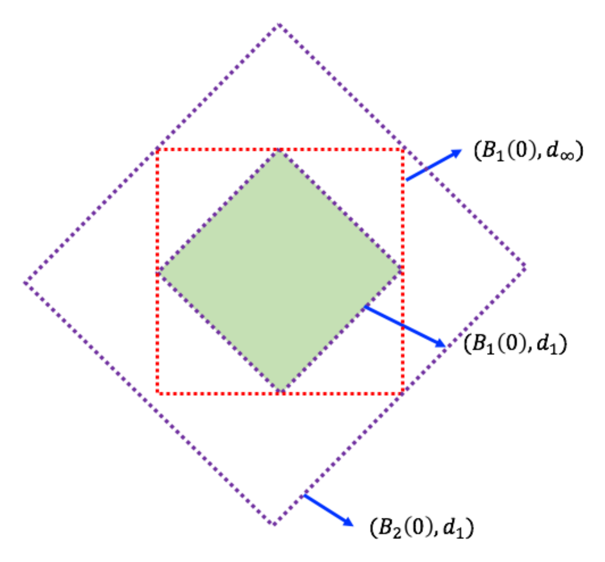


Figure 1.5: The diagram for the relation $(B_1(x), d_1) \subseteq (B_\infty(x), d_\infty) \subseteq (B_2(x), d_1)$ on $\mathbb{R}^2$

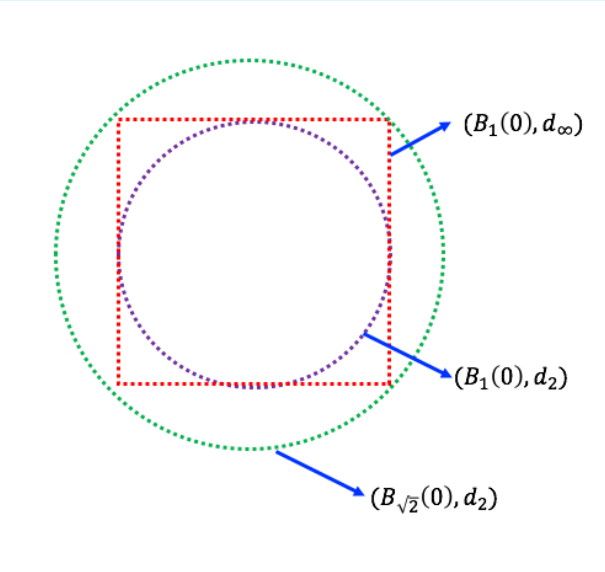


Figure 1.6: The diagram for the relation $(B_1(x), d_2) \subseteq (B_\infty(x), d_\infty) \subseteq (B_{\sqrt{2}}(x), d_2)$ on $\mathbb{R}^2$

It's easy to conclude the simple generalization for example (1.16):

Proposition 1.11 If d and ρ are equivalent, then

$$\lim_{n\to\infty} d(x_n, x) = 0 \Longleftrightarrow \lim_{n\to\infty} \rho(x_n, x) = 0$$

Note that this does not necessarily hold for topology spaces.

2. Consider d_1, d_∞ in $C[a,b]$:

$$d_1(f,g) := \int_a^b |f-g| \, \mathrm{d}x \le \int_a^b \sup_{[a,b]} |f-g| \, \mathrm{d}x = (b-a) d_\infty(f,g),$$

i.e., d_∞ is stronger than d_1. Question: Are they equivalent? **No**.

Justification. Consider $f_n(x) = n^2 x^n (1-x)$ for $x \in [0,1]$. Check that

$$\lim_{n\to\infty} d_1(f_n(x), 0) = 0, \quad \text{but } d_\infty(f_n(x), 0) \to \infty$$

The peak of f_n may go to infinite, while the integration converges to zero, i.e., there is no $K > 0$ such that $d_\infty(f_n, 0) < K d_1(f_n, 0), \forall n \in \mathbb{N}$. ■

We will discuss this topic at Lebsegue integration again.

1.5.2. Continuity

Definition 1.16 [Continuity] Let $f : (X,d) \to (Y,d)$ be a function and $x_0 \in X$. Then f is continuous at x_0 if $\forall \varepsilon > 0$, there exists $\delta > 0$ such that

$$d(x, x_0) < \delta \implies \rho(f(x), f(x_0)) < \varepsilon$$

The function f is continuous in X if f is continuous for all $x_0 \in X$.

Proposition 1.12 The function f is continuous at x if and only if for all $\{x_n\} \to x$ under d, $f(x_n) \to f(x)$ under ρ.

Proof. *Necessity:* Given $\varepsilon > 0$, by continuity,

$$d(x, x') < \delta \implies \rho(f(x'), f(x)) < \varepsilon. \tag{1.3}$$

Consider the sequence $\{x_n\} \to x$, then there exists N such that $d(x_n, x) < \delta$ for $\forall n \ge N$. By applying (1.3), $\rho(f(x_n), f(x)) < \varepsilon$ for $\forall n \ge N$, i.e., $f(x_n) \to f(x)$.

Sufficiency: Assume that f is not continuous at x, then there exists ε_0 such that for

$\delta_n = \frac{1}{n}$, there exists x_n such that

$$d(x_n, x) < \delta_n, \text{ but } \rho(f(x_n), f(x)) > \varepsilon_0.$$

Then $\{x_n\} \to x$ by our construction, while $\{f(x_n)\}$ does not converge to $f(x)$, which is a contradiction. ■

Corollary 1.2 If the function $f : (X, d) \to (Y, \rho)$ is continuous at x, the function $g : (Y, \rho) \to (Z, m)$ is continuous at $f(x)$, then $g \circ f : (X, d) \to (Z, m)$ is continuous at x.

Proof. Note that

$$\{x_n\} \to x \overset{(a)}{\Longrightarrow} \{f(x_n)\} \to f(x) \overset{(b)}{\Longrightarrow} \{g(f(x_n))\} \to g(f(x)) \overset{(c)}{\Longrightarrow} g \circ f \text{ is continuous at } x.$$

where $(a), (b), (c)$ are all by proposition (1.12). ■

1.5.3. Open and Closed Sets

We have open/closed intervals in $\mathbb{R}$, and they are important in some theorems (e.g, continuous functions bring closed intervals to closed intervals).

Definition 1.17 [Open] Let (X, d) be a metric space. A set $U \subseteq X$ is open if for each $x \in U$, there exists $\rho_x > 0$ such that $B_{\rho_x}(x) \subseteq U$. The empty set $\emptyset$ is defined to be open. ▪

▪ **Example 1.18** Let $(\mathbb{R}, d_2 \text{ or } d_\infty)$ be a metric space. The set $U = (a, b)$ is open. ▪

Proposition 1.13
1. Let (X, d) be a metric space. Then all open balls $B_r(x)$ are open
2. All open sets in X can be written as a union of open balls.

Proof.
1. Let $y \in B_r(x)$, i.e., $d(x, y) := q < r$. Consider the open ball $B_{(r-q)/2}(y)$. It suffices to show $B_{(r-q)/2}(y) \subseteq B_r(x)$. For any $z \in B_{(r-q)/2}(y)$,

$$d(x, z) \le d(x, y) + d(y, z) < q + \frac{r - q}{2} = \frac{r + q}{2} < r.$$

The proof is complete.

2. Let $U \subseteq X$ be open, i.e., for $\forall x \in U$, there exists $\varepsilon_x > 0$ such that $B_{\varepsilon_x}(x) \subseteq U$. Therefore

$$\{x\} \subseteq B_{\varepsilon_x}(x) \subseteq U, \forall x \in U$$

which implies

$$U = \bigcup_{x \in U} \{x\} \subseteq \bigcup_{x \in U} B_{\varepsilon_x}(x) \subseteq U,$$

i.e., $U = \bigcup_{x \in U} B_{\varepsilon_x}(x)$.

■

2.2. Monday for MAT3006

Reviewing.

1. Continuous functions: the function f is continuous is equivalent to say for $\forall x_n \to x$, we have $f(x_n) \to f(x)$.
2. Open sets: Let (X,d) be a metric space. A set $U \subseteq X$ is open if for each $x \in U$, there exists $\rho_x > 0$ such that $B_{\rho_x}(x) \subseteq U$.

R Unless stated otherwise, we assume that

$$C[a,b] \longleftrightarrow (C[a,b], d_\infty)$$

$$\mathbb{R}^n \longleftrightarrow (\mathbb{R}^n, d_2)$$

2.2.1. Remark on Open and Closed Set

■ **Example 2.6** Let $X = C[a,b]$, show that the set

$$U := \{f \in X \mid f(x) > 0, \forall x \in [a,b]\} \quad \text{is open.}$$

Take a point $f \in U$, then

$$\inf_{[a,b]} f(x) = m > 0.$$

Consider the ball $B_{m/2}(f)$, and for $\forall g \in B_{m/2}(f)$,

$$\begin{aligned}
|g(x)| &\geq |f(x)| - |f(x) - g(x)| \\
&\geq \inf_{[a,b]} |f(x)| - \sup_{[a,b]} |f(x) - g(x)| \\
&\geq m - \frac{m}{2} \\
&= \frac{m}{2} > 0, \ \forall x \in [a,b]
\end{aligned}$$

Therefore, we imply $g \in U$, i.e., $B_{m/2}(f) \subseteq U$, i.e., U is open in X. ∎

Proposition 2.2 Let (X,d) be a metric space. Then

1. $\emptyset, X$ are open in X
2. If $\{U_\alpha \mid \alpha \in \mathcal{A}\}$ are open in X, then $\bigcup_{\alpha \in \mathcal{A}}$ is also open in X
3. If $U_1, \dots, U_n$ are open in X, then $\bigcap_{i=1}^n U_i$ are open in X

(R) Note that $\bigcap_{i=1}^\infty U_i$ is not necessarily open if all U_i's are all open:

$$\bigcap_{i=1}^\infty \left(-\frac{1}{i}, 1 + \frac{1}{i}\right) = [0,1]$$

Definition 2.2 [Closed] The closed set in metric space (X,d) are the complement of open sets in X, i.e., any closed set in X is of the form $V = X \setminus U$, where U is oepn. ∎

For example, in $\mathbb{R}$,

$$[a,b] = \mathbb{R} \setminus \{(-\infty, a) \bigcup (b, \infty)\}$$

Proposition 2.3

1. $\emptyset, X$ are closed in X
2. If $\{V_\alpha \mid \alpha \in \mathcal{A}\}$ are closed subsets in X, then $\bigcap_{\alpha \in \mathcal{A}} V_\alpha$ is also closed in X
3. If $V_1, \dots, V_n$ are closed in X, then $\bigcup_{i=1}^n V_i$ is also closed in X.

(R) Whenever you say U is open or V is closed, you need to specify the underlying space, e.g.,

$$\textbf{Wrong}: U \text{ is open}$$

$$\textbf{Right}: U \text{ is open in } X$$

Proposition 2.4 The following two statements are equivalent:

1. The set V is closed in metric space (X,d).
2. If the sequence $\{v_n\}$ in V converges to x, then $x \in V$

Proof. *Necessity.*

Suppose on the contrary that $\{v_n\} \to x \notin V$. Since $X \setminus V \ni x$ is open, there exists an open ball $B_\varepsilon(x) \subseteq X \setminus V$.

Due to the convergence of sequence, there exists N such that $d(v_n, x) < \varepsilon$ for $\forall n \geq N$, i.e., $v_n \in B_\varepsilon(x)$, i.e., $v_n \notin V$, which contradicts to $\{v_n\} \subseteq V$.

Sufficiency.

Suppose on the contrary that V is not closed in X, i.e., $X \setminus V$ is not open, i.e., there exists $x \notin V$ such that for all open $U \ni x$, $U \bigcap V \neq \emptyset$. In particular, take

$$U_n = B_{1/n}(x), \implies \exists v_n \in B_{1/n}(x) \bigcap V,$$

i.e., $\{v_n\} \to x$ but $x \notin V$, which is a contradiction. ∎

Proposition 2.5 Given two metric space (X,d) and (Y,ρ), the following statements are equivalent:

1. A function $f : (X,d) \to (Y,\rho)$ is continuous on X
2. For $\forall U \subseteq Y$ open in Y, $f^{-1}(U)$ is open in X.
3. For $\forall V \subseteq Y$ closed in Y, $f^{-1}(V)$ is closed in X.

▪ **Example 2.7** The mapping $\Psi : C[a,b] \to \mathbb{R}$ is defined as:

$$f \mapsto f(c)$$

where Ψ is called a **functional**.

Show that Ψ is continuous by using d_∞ metric on $C[a,b]$:

1. Any open set in $\mathbb{R}$ can be written as countably union of open disjoint intervals, and therefore suffices to consider the pre-image $\Psi^{-1}(a,b) = \{f \mid f(c) \in (a,b)\}$. Following the similar idea in Example (2.6), it is clear that $\Psi^{-1}(a,b)$ is open in $(C[a,b], d_\infty)$. Therefore, Ψ is continuous.
2. Another way is to apply definition.

▪

We now study open sets in a subspace $(Y, d_Y) \subseteq (X, d_X)$, i.e.,

$$d_Y(y_1, y_2) := d_X(y_1, y_2), \quad \forall y_1, y_2 \in Y.$$

Therefore, the open ball is defined as

$$\begin{aligned} B_\varepsilon^Y(y) &= \{y' \in Y \mid d_Y(y, y') < \varepsilon\} \\ &= \{y' \in Y \mid d_X(y, y') < \varepsilon\} \\ &= \{y' \in X \mid d_X(y, y') < \varepsilon, y' \in Y\} \\ &= B_\varepsilon^X(y) \bigcap Y \end{aligned}$$

Proposition 2.6 All open sets in the subspace $(Y, d_Y) \subseteq (X, d_X)$ are of the form $U \bigcap Y$, where U is open in X.

Corollary 2.1 For the subspace $(Y, d_Y) \subseteq (X, d_X)$, the mapping $i : (Y, d_Y) \to (X, d_X)$ with $i(y) = y, \forall y \in Y$ is continuous.

Proof. $i^{-1}(U) = U \bigcap Y$ for any subset $U \subseteq X$. The results follows from proposition (2.5). ∎

(R) It's important to specify the underlying space to describe an open set.

For example, the interval $[0, \frac{1}{2})$ is not open in $\mathbb{R}$, while $[0, \frac{1}{2})$ is open in $[0, 1]$, since

$$[0, \frac{1}{2}) = (-\frac{1}{2}, \frac{1}{2}) \bigcap [0, 1].$$

2.2.2. Boundary, Closure, and Interior

Definition 2.3 Let (X, d) be a metric space, then

1. A point x is a **boundary point** of $S \subseteq X$ (denoted as $x \in \partial S$) if for any open $U \ni x$, then both $U \bigcap S, U \setminus S$ are non-empty.

(one can replace U by $B_{1/n}(x)$, with $n = 1,2,\ldots$)

2. The **closure** of S is defined as $\overline{S} = S \bigcup \partial S$.
3. A point x is an **interior point** of S (denoted as $x \in S^\circ$) if there $\exists U \ni x$ open such that $U \subseteq S$. We use S° to denote the set of interior points.

Proposition 2.7 1. The closure of S can be equivalently defined as

$$\overline{S} = \bigcap \{C \in X \mid C \text{ is closed and } C \supseteq S\}$$

Therefore, $\overline{S}$ is the smallest closed set containing S.

(Note that $C = X$ is a closed set containing S, and hence $\overline{S}$ is well-defined.)

2. The interior set of S can be equivalently defined as

$$S^\circ = \bigcup \{U \subseteq X \mid U \text{ is open and } U \subseteq S\}$$

Therefore, S° is the largest open set contained in S.

Example 2.8 For $S = [0, \frac{1}{2}] \subseteq X$, we have

1. $\partial S = \{0, \frac{1}{2}\}$
2. $\overline{S} = [0, \frac{1}{2}]$
3. $S^\circ = (0, \frac{1}{2})$

Proof. 1. (a) Firstly, we show that $\overline{S}$ is closed, i.e., $X \setminus \overline{S}$ is open.

- Take $x \notin \overline{S}$. Since $x \notin \partial S$, there $\exists B_r(x) \ni x$ such that

$$B_r(x) \bigcap S, \quad \text{or} \quad B_r(x) \setminus S \text{ is } \emptyset.$$

- Since $x \notin S$, the set $B_r(x) \setminus S$ is not empty. Therefore, $B_r(x) \bigcap S = \emptyset$.
- It's clear that $B_{r/2}(x) \bigcap S = \emptyset$. We claim that $B_{r/2}(x) \bigcap \overline{S}$ is empty.

Suppose on the contrary that

$$y \in B_{r/2}(x) \bigcap \partial S,$$

which implies that $B_{r/2}(y) \bigcap S \neq \emptyset$. Therefore,

$$B_{r/2}(y) \subseteq B_r(x) \implies B_r(x) \bigcap S \supseteq B_{r/2}(y) \bigcap S \neq \emptyset,$$

which is a contradiction.

Therefore, $x \in X \setminus \overline{S}$ implies $B_{r/2}(x) \bigcap \overline{S} = \emptyset$, i.e., $X \setminus \overline{S}$ is open, i.e., $\overline{S}$ is closed.

(b) Secondly, we show that $\overline{S} \subseteq C$, for any closed $C \supseteq S$, i.e., suffices to show $\partial S \subseteq C$.

Take $x \in \partial S$, which implies that $B_\varepsilon(x) \cap S$ is non-empty for any $\varepsilon > 0$. Therefore, construct a sequence

$$x_n \in B_{1/n}(x) \bigcap S.$$

Here $\{x_n\}$ is a sequence in $S \subseteq C$ converging to x, which implies $x \in C$, due to the closeness of C in X.

Combining (a) and (b), the result follows naturally.

2. Exercise. Show that

$$S^\circ = S \setminus \partial S = X \setminus (\overline{X \setminus S}).$$

Then it's clear that S° is open, and contained in S.

■

The next lecture we will talk about compactness and sequential compactness.

2.5. Wednesday for MAT3006

2.5.1. Compactness

This lecture will talk about the generalization of closeness and boundedness property in $\mathbb{R}^n$. First let's review some simple definitions:

Definition 2.11 [Compact] Let (X,d) be a metric space, and $\{U_\alpha\}_{\alpha\in\mathcal{A}}$ a collection of open sets.

1. $\{U_\alpha\}_{\alpha\in\mathcal{A}}$ is called an **open cover** of $E\subseteq X$ if $E\subseteq\cup_{\alpha\in\mathcal{A}}U_\alpha$
2. A **finite subcover** of $\{U_\alpha\}_{\alpha\in\mathcal{A}}$ is a finite sub-collection $\{U_{\alpha_1},\dots,U_{\alpha_n}\}\subseteq\{U_\alpha\}$ covering E.
3. The set $E\subseteq X$ is **compact** if every open cover of E has a finite subcover.

A well-known result is talked in MAT2006:

Theorem 2.4 — Heine-Borel Theorem. The set $E\subseteq\mathbb{R}^n$ is **compact** if and only if E is closed and bounded.

However, there's a notion of sequentially compact, and we haven't identify its gap and relation with compactness.

Definition 2.12 [Sequentially Compact] Let (X,d) be a metric space. Then $E\subseteq X$ is **sequentially compact** if every sequence in E has a convergent subsequence with limit in E.

A well-known result is talked in MAT2006:

Theorem 2.5 — Bolzano-Weierstrass Theorem. The set $E\subseteq\mathbb{R}^n$ is closed and bounded if and only if E is sequentially compact.

Actually, the definitions of comapctness and the sequential compactness are equivalent under a metric space.

Theorem 2.6 Let (X,d) be a metric space, then $E \subseteq X$ is compact if and only if E is sequentially compact.

Proof. *Necessity*

Suppose $\{x_n\}$ is a sequence in E, it suffices to show it has a convergent subsequence. Consider the tail of $\{x_n\}$, say

$$F_n = \overline{\{x_k \mid k \ge n\}} \implies F_1 \supseteq F_2 \supseteq \cdots.$$

- Note that $\cap_{i=1}^{\infty} F_i \neq \emptyset$. Assume not, then we imply $\cup_{i=1}^{\infty}(E \setminus F_i) = E$, i.e., $\{E \setminus F_i\}_{i=1}^{\infty}$ a open cover of E. By the compactness of E, we imply there exists a finite subcover of E:
$$E = \bigcup_{j=1}^{r}(E \setminus F_{i_j}) \implies \bigcap_{j=1}^{r} F_{i_j} = \emptyset \implies F_{i_r} = \emptyset,$$
which is a contradiction, and there must exist an element $x \in \cap_{n=1}^{\infty} F_i$.

- For any $n \ge 1$ and $x \in \cap_{n=1}^{\infty} F_i$, either $x \in \{x_k \mid k \ge n\}$ or $x \in \partial\{x_k \mid k \ge n\}$. In both cases, the open ball $B_\varepsilon(x)$ must intersect with the n-th tail of the sequence $\{x_n\}$ for any $\varepsilon > 0$:
$$B_\varepsilon(x) \cap \{x_k \mid k \ge n\} \neq \emptyset, \ \forall \varepsilon > 0.$$
Therefore, construct $x_{n_1} \in B_1(x) \cap \{x_k \mid k \ge 1\}$ and for $r > 1$,
$$x_{n_r} \in B_{1/r}(x) \cap \{x_k \mid k \ge n_{r-1} + 1\}.$$
Therefore, the subsequence $x_{n_r} \to x$ as $r \to \infty$. The proof for necessity is complete.

Sufficiency

Firstly, let's assume the claim below hold (which will be shown later):

Proposition 2.19 If $E \subseteq X$ is sequentially compact, then for any$\varepsilon > 0$, there exists finitely many open balls, say $\{B_\varepsilon(x_1), \dots, B_\varepsilon(x_n)\}$, covering E.

Suppose on the contrary that there exists an open cover $\{U_\alpha\}_{\alpha \in \mathcal{A}}$ of E, that has no finite subcover.

- By proposition (2.19), for $n \geq 1$, there are finitely many balls of radius $1/n$ covering E. Due to our assumption, there exists a open ball $B_{1/n}(y_n)$ such that $B_{1/n}(y_n) \cap E$ cannot be covered by finitely many members in $\{U_\alpha\}_{\alpha \in \mathcal{A}}$.

- Pick $x_n \in B_{1/n}(y_n)$ to form a sequence. Due to the sequential compactness of E, there exists a subsequence $\{x_{n_j}\} \to x$ for some $x \in E$.

- Since $\{U_\alpha\}_{\alpha \in \mathcal{A}}$ covers E, there exists a U_β containing x. Since U_β is open and the radius of $B_{1/n_j}(y_{n_j})$ tends to 0, we imply that, for sufficiently large n_j, the set $B_{1/n_j}(y_{n_j}) \cap E$ is contained in U_β.

  In oteher words, U_β forms a **single** subcover of $B_{1/n}(y) \cap E$, which contradicts to our choice of $B_{1/n_j}(y_{n_j}) \cap E$. The proof for sufficiency is complete.

■

Proof for proposition (2.19). Pick $B_\varepsilon(x_1)$ for some $x_1 \in E$. Suppose $E \setminus B_\varepsilon(x_1) \neq \emptyset$. We can find $x_2 \notin B_\varepsilon(x_1)$ such that $d(x_2, x_1) \geq \varepsilon$.

Suppose $E \setminus (B_\varepsilon(x_1) \bigcup B_\varepsilon(x_2))$ is non-empty, then we can find $x_3 \notin B_\varepsilon(x_1) \bigcup B_\varepsilon(x_2)$ so that $d(x_j, x_3) \geq \varepsilon$, $j = 1, 2$.

Keeping this procedure, we obtain a sequence $\{x_n\}$ in E such that

$$E \setminus \bigcup_{j=1}^{n} B_\varepsilon(x_j) \neq \emptyset, \qquad \text{and} \qquad d(x_j, x_n) \geq \varepsilon, j = 1, 2, \ldots, n-1.$$

By the sequential compactness of E, there exists $\{x_{n_j}\}$ and $x \in E$ so that $x_{n_j} \to x$ as $j \to \infty$. But then $d(x_{n_j}, x_{n_k}) < d(x_{n_j}, x) + d(x_{n_k}, x) \to 0$, which contradicts that $d(x_j, x_n) \geq \varepsilon$ for $\forall j < n$.

Therefore, one must have $E \setminus \bigcup_{j=1}^{N} B_\varepsilon(x_j) = \emptyset$ for some finite N.

The proof is complete. ■

R

1. Given the condition metric space,

$$\text{Sequential Compactness} \Longleftrightarrow \text{Compactness}$$

2. Given the condition metric space, we will show that

$$\text{Compactness} \implies \text{Closed and Bounded}$$

However, the converse may not necessarily hold. Given the condition the metric space is $\mathbb{R}^n$, then

$$\text{Compactness} \iff \text{Closed and Bounded}$$

Proposition 2.20 Let (X,d) be a metric space. Then $E \subseteq X$ is compact implies that E is closed and bounded.

We say a set E if bounded if there exists $K \geq 0$ such that

$$d(e_1,e_2) < K, \quad \forall e_1,e_2 \in E$$

Proof. 1. Let $\{x_n\}$ be a convergent sequence in E. By sequential compactness, $\{x_{n_j}\} \to x$ for some $x \in E$. By the uniqueness of limits, under metric space, $\{x_n\} \to x$ for $x \in E$. The closeness is shown

2. Take $x \in E$ and consider the open cover $\bigcup_{n=1}^{\infty} B_n(x)$ of E. By compactness,

$$E \subseteq \bigcup_{i=1}^{k} B_{n_i}(x) = B_{n_k}(x),$$

which implies that for any $y,z \in E$,

$$d(y,z) \leq d(y,x) + d(x,z) \leq n_k + n_k = 2n_k.$$

The boundness is shown.

∎

Here we raise several examples to show that the coverse does not necessarily hold under a metric space.

■ **Example 2.14** Given the metric space $C[0,1]$ and a set $E = \{f \in C[0,1] \mid 0 \le f(x) \le 1\}$.

Notice that E is closed and bounded:

- $E = \cap_{x\in[0,1]}\Psi_x^{-1}([0,1])$, where $\Psi_x(f) = f(x)$, which implies that E is closed.
- Note that $E \subseteq B_2(\mathbf{0}) = \{f \mid |f| < 2\}$, i.e., E is bounded.

However, E may not be compact. Consider a sequence $\{f_n\}$ with

$$f_n(x) = \begin{cases} nx, & 0 \le x \le \dfrac{1}{n} \\ 1, & \dfrac{1}{n} \le x \le 1 \end{cases}$$

Suppose on the contrary that E is sequentially compact, therefore there exists a subsequence $\{f_{n_k}\} \to f$ under d_∞ metric, which implies, $\{f_{n_k}\}$ uniformly converges to f. By the definition of $f_n(x)$, we imply

$$f(x) = \begin{cases} 0, & x = 0 \\ 1, & x \in (0,1] \end{cases}$$

However, since d_∞ indicates uniform convergence, the limit for $\{f_{n_k}\}$, say f, must be continuous, which is a contradiction.

■

Theorem 2.7 Let the set E be compact in (X,d) and the function $f : (X,d) \to (Y,\rho)$ is continuous. Then $f(E)$ is compact in Y.

Note that the technique to show compactness by using the sequential compactness is very useful. However, this technique only applies to the metric space, but fail in general topological spaces.

Proof. Let $\{y_n\} = \{f(x_n)\}$ be any sequence in $f(E)$.

- By the compactness of X, $\{x_n\}$ has a convergent subsequence $\{x_{n_r}\} \to x$ as $r \to \infty$.
- Therefore, $\{y_{n_r}\} := \{f(x_{n_r})\} \to f(x)$ by the continuity of f.

- Therefore, $f(E)$ is sequentially compact, i.e., compact.

∎

(R) The Theorem (2.7) is a generalization of the statement that *a continuous function on $\mathbb{R}^n$ admits its minimum and maximum*. Note that such an extreme value property no longer holds for arbitrary closed, bounded sets in a general metric space, but it continues to hold when the sets are strengthened to compact ones.

Another characterization of compactness in $C[a,b]$ is shown in the Ascoli-Arzela Theorem (see Theorem (14.1) in MAT2006 Notebook).

2.5.2. Completeness

Definition 2.13 [Complete] Let (X,d) be metric space.

1. A sequence $\{x_n\}$ in (X,d) is a **Cauchy sequence** if for every $\varepsilon > 0$, there exists some N such that $d(x_n,x_m) < \varepsilon$ for all $n,m \ge N$.
2. A subset $E \subseteq X$ is said to be **complete** if every Cauchy sequence in E is convergent.

■ **Example 2.15** The set $X = C[a,b]$ is complete:

- Suppose $\{f_n\}$ is Cauchy in $C[a,b]$, i.e., $\{f_n(x)\}$ is Cauchy in $\mathbb{R}$ for $\forall x \in [a,b]$.
- By the completeness of $\mathbb{R}$, the sequence $f_n(x) \to f(x)$ for some $f(x) \in \mathbb{R}, \forall x \in [a,b]$. It suffices to show $f_n \to f$ uniformly:
  - For fixed $\varepsilon > 0$, there exists $N > 0$ such that

$$d_\infty(f_n, f_{n+k}) < \frac{\varepsilon}{2}, \qquad \forall n \ge N, k \in \mathbb{N}$$

which implies that for $\forall x \in [a, b]$, $\forall n \geq N, k \in \mathbb{N}$,

$$|f_n(x) - f_{n+k}(x)| < \frac{\varepsilon}{2} \implies \lim_{k \to \infty} |f_n(x) - f_{n+k}(x)| \leq \frac{\varepsilon}{2}$$

Therefore, we imply

$$|f_n(x) - f(x)| = \lim_{k \to \infty} |f_n(x) - f_{n+k}(x)| \leq \frac{\varepsilon}{2} < \varepsilon, \qquad \forall n \geq N, x \in [a, b]$$

The proof is complete.

3.2. Monday for MAT3006

Reviewing.

1. Compactness/Sequential Compactness:
   - Equivalence for metric space
   - Stronger than closed and bounded
2. Completeness:
   - The metric space (E,d) is complete if every Cauchy sequence on E is convergent.
   - $\mathbb{P}[a,b] \subseteq C[a,b]$ is not complete:

$$f_N(x) = \sum_{n=0}^{N}(-1)^n \frac{x^{2n}}{(2n)!} \to \cos x,$$

     while $\cos x \notin \mathcal{P}[a,b]$.

3.2.1. Remarks on Completeness

Proposition 3.3 Let (X,d) be a metric space.

1. If X is complete and $E \subseteq X$ is closed, then E is complete.
2. If $E \subseteq X$ is complete, then E is closed in X.
3. If $E \subseteq X$ is compact, then E is complete.

Proof.
1. Every Cauchy sequence $\{e_n\}$ in $E \subseteq X$ is also a Cauchy sequence in E.
   Therefore we imply $\{e_n\} \to x \in X$, due to the completeness of X.
   Due to the closedness of E, the limit $x \in E$, i.e., E is complete.
2. Consider any convergent sequence $\{e_n\}$ in E, with some limit $x \in X$.
   We imply $\{e_n\}$ is Cauchy and thus $\{e_n\} \to e \in E$, due to the completeness of E.
   By the uniqueness of limits, we must have $x = z \in E$, i.e., E is closed.
3. Consider a Cauchy sequence $\{e_n\}$ in E. There exists a subsequence $\{e_{n_j}\} \to e \in E$, due to the sequential compactness of E.

It follows that for large n and j,

$$d(e_n, e) \overset{(a)}{\leq} d(e_n, e_{n_j}) + d(e_{n_j}, e) \overset{(b)}{<} \varepsilon$$

where (a) is due to triangle inequality and (b) is due to the Cauchy property of $\{e_n\}$ and the convergence of $\{e_{n_j}\}$.

Therefore, we imply $\{e_n\} \to e \in E$, i.e., E is complete.

∎

R Given any metric space that may not be necessarily complete, we can make the union of it with another space to make it complete, e.g., just like the completion from $\mathbb{Q}$ to $\mathbb{R}$.

3.2.2. Contraction Mapping Theorem

The motivation of the contraction mapping theorem comes from solving an equation $f(x)$. More precisely, such a problem can be turned into a problem for fixed points, i.e., it suffices to find the fixed points for $g(x)$, with $g(x) = f(x) + x$.

Definition 3.3 Let (X, d) be a metric space. A map $T : (X, d) \to (X, d)$ is a **contraction** if there exists a constant $\tau \in (0, 1)$ such that

$$d(T(x), T(y)) < \tau \cdot d(x, y), \quad \forall x, y \in X$$

A point x is called a fixed point of T if $T(x) = x$.

R All contractions are continuous: Given any convergence sequence $\{x_n\} \to x$, for $\varepsilon > 0$, take N such that $d(x_n, x) < \frac{\varepsilon}{\tau}$ for $n \geq N$. It suffices to show the convergence of $\{T(x_n)\}$:

$$d(T(x_n), T(x)) < \tau \cdot T(x_n, x) < \tau \cdot \frac{\varepsilon}{\tau} = \varepsilon.$$

Therefore, the contraction is Lipschitz continuous with Lipschitz constant τ.

Theorem 3.2 — Contraction Mapping Theorem / Banach Fixed Point Theorem. Every contraction T in a **complete** metric space X has a unique fixed point.

■ **Example 3.5** 1. The mapping $f(x) = x + 1$ is not a contraction in $X = \mathbb{R}$, and it has no fixed point.

2. Consider an in-complete space $X = (0,1)$ and a contraction $f(x) = \frac{x+1}{2}$. It doesn't admit a fixed point on X as well.

■

Proof. Pick any $x_0 \in X$, and define a sequence recursively by setting $x_{n+1} = T(x_n)$ for $n \geq 0$.

1. Firstly show that the sequence $\{x_n\}$ is Cauchy.

   We can upper bound the term $d(T^n(x_0), T^{n-1}(x_0))$:

$$d(T^n(x_0), T^{n-1}(x_0)) \leq \tau d(T^{n-1}(x_0), T^{n-2}(x_0)) \leq \cdots \leq \tau^{n-1} d(T(x_0), x_0) \tag{3.4}$$

   Therefore for any $n \geq m$, where m is going to be specified later,

$$\begin{aligned}
d(x_n, x_m) &= d(T^n(x_0), T^m(x_0)) && (3.5a)\\
&\leq \tau d(T^{n-1}(x_0), T^{m-1}(x_0)) \leq \cdots \leq \tau^m d(T^{n-m}(x_0), x_0) && (3.5b)\\
&\leq \tau^m \sum_{j=1}^{n-m} \tau^{n-m-j} d(T(x_0), x_0) && (3.5c)\\
&< \frac{\tau^m}{1-\tau} d(T(x_0), x_0) && (3.5d)\\
&\leq \varepsilon && (3.5e)
\end{aligned}$$

   where (3.5b) is by repeatedly applying contraction property of d; (3.5c) is by applying the triangle inequality and (3.4); (3.5e) is by choosing sufficiently large m such that $\frac{\tau^m}{1-\tau} d(T(x_0), x_0) < \varepsilon$.

   Therefore, $\{x_n\}$ is Cauchy. By the completeness of X, we imply $\{x_n\} \to x \in X$.

2. Therefore, we imply

$$x = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} T(x_n) = T(\lim_{n\to\infty} x_n) = T(x),$$

i.e., x is a fixed point of T.

Now we show the uniqueness of the fixed point. Suppose that there is another fixed point $y \in X$, then

$$d(x,y) = d(T(x),T(y)) < \tau \cdot d(x,y) \implies d(x,y) < \tau d(x,y), \quad \tau \in (0,1),$$

and we conclude that $d(x,y) = 0$, i.e., $x = y$.

■

■ **Example 3.6** [Convergence of Newton's Method] The Newton's method aims to find the root of $f(x)$ by applying the iteration

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

Suppose r is a root for f, the pre-assumption for the convergence of Newton's method is:

1. $f'(r) \neq 0$
2. $f \in C^2$ on some neighborhood of r

Proof. 1. We first show that there exists $[r-\varepsilon, r+\varepsilon]$ such that the mapping

$$\begin{aligned} T: \quad & \mathbb{R} \to \mathbb{R}, \\ \text{with} \quad & x \mapsto x - \tfrac{f(x)}{f'(x)} \end{aligned}$$

satisfies

$$|T'(x)| < 1, \quad \forall x \in [r-\varepsilon, r+\varepsilon]. \tag{3.6}$$

Note that $T'(x) = \frac{f(x)}{[f'(x)]^2} f''(x)$, and we define $h(x) = |T'(x)|$.

It's clear that $h(r) = 0$ and $h(x)$ is continuous, which implies

$$r \in h^{-1}((-1,1)) \implies B_\rho(r) \subseteq h^{-1}((-1,1)) \text{ for some } \rho > 0.$$

Or equivalently, $h((r-\rho, r+\rho)) \subseteq (-1,1)$. Take $\varepsilon = \frac{\rho}{2}$, and the result is obvious.

2. Moreover, for any $x, y \in [r-\varepsilon, r+\varepsilon]$,

$$\begin{aligned} d(T(x), T(y)) :&= |T(x) - T(y)| && (3.7a) \\ &= |T'(\xi)||x-y| && (3.7b) \\ &\le \max_{\xi \in [r-\varepsilon, r+\varepsilon]} |T'(\xi)||x-y| && (3.7c) \\ &< m \cdot |x-y| && (3.7d) \end{aligned}$$

where (3.7b) is by applying MVT, and ξ is some point in $[r-\varepsilon, r+\varepsilon]$; we assume that $\max_{\xi \in [r-\varepsilon, r+\varepsilon]} |T'(\xi)| < m$ for some $m < 1$ in (3.7d).

3. Note that (2) is not enough to show that T is a contraction. We further need to show that $T(x) \in [r-\varepsilon, r+\varepsilon]$ provided that $x \in [r-\varepsilon, r+\varepsilon]$:

$$|T(x) - r| = |T(x) - T(r)| = |T'(s)||x-r| \le \sup_{[r-\varepsilon, r+\varepsilon]} |T'(s)||x-r| < |x-r|.$$

Combining (2) and (3), we imply T is a contraction on $[r-\varepsilon, r+\varepsilon]$. By applying the contraction mapping theorem, there exists a unique fixed point near $[r-\varepsilon, r+\varepsilon]$:

$$x - \frac{f(x)}{f'(x)} = x \implies \frac{f(x)}{f'(x)} = 0 \implies f(x) = 0,$$

i.e., we obtain a root $x = r$. ■

Summary: if we use Newton's method on any point between $[r-\varepsilon, r+\varepsilon]$ where $f(r) = 0$ and ε is sufficiently small, then we will eventually get close to r.

3.2.3. Picard Lindelof Theorem

We will use Banach fixed point theorem to show the existence and uniqueness of the solution of ODE

$$\begin{cases} \dfrac{\mathrm{d}y}{\mathrm{d}x} = f(x, y(x)) \\ y(x_0) = y_0 \end{cases} \qquad \textbf{Initial Value Problem, IVP} \tag{3.8}$$

■ **Example 3.7** Consider the IVP

$$\begin{cases} \dfrac{\mathrm{d}y}{\mathrm{d}x} = x^2 y^{1/5} \\ y(x_0) = c > 0 \end{cases} \implies y = \left(\frac{4x^3}{15} + c^{4/5}\right)^{5/4}$$

which can be solved by the separation of variables:

$$c > 0 \implies y = \left(\frac{4x^3}{15} + c^{4/5}\right)^{5/4}.$$

However, when $c = 0$, the ODE does not have a unique solution. One can verify that y_1, y_2 given below are both solutions of this ODE:

$$y_1 = (\frac{4x^3}{15})^{5/4}, \qquad y_2 = 0$$

This example shows that even when f is very nice, the IVP may not have unique solution. The Picard-Lindelof theorem will give a clean condition on f ensuring the unique solvability of the IVP (3.8). ■

3.5. Wednesday for MAT3006

3.5.1. Remarks on Contraction

Reviewing.

- Suppose $E \subseteq X$ with X being complete, then E is closed in X iff E is complete
- Suppose $E \subseteq X$, then E is closed in X if E is complete.
- Contraction Mapping Theorem

3.5.2. Picard-Lindelof Theorem

Consider solving the the initival value problem given below

$$\begin{cases} \frac{\mathrm{d}y}{\mathrm{d}x} = f(x,y) \\ y(\alpha) = \beta \end{cases} \implies y(x) = \beta + \int_{\alpha}^{x} f(t, y(t))\,\mathrm{d}t \tag{3.11}$$

Definition 3.5 Let $R = [\alpha - a, \alpha + a] \times [\beta - b, \beta + b]$. Then the function $f(x,y)$ satisfies the **Lipschitz condition** on R if there exists $L > 0$ such that

$$|f(x,y_1) - f(x,y_2)| < L \cdot |y_1 - y_2|, \quad \forall (x, y_i) \in R \tag{3.12}$$

The smallest number $L^* = \inf\{L \mid \text{The relation (3.12) holds for } L\}$ is called the **Lipschitz constant** for f.

Example 3.11 A C^1-function f(x,y) in a rectangle automatically satisfies the Lipschitz condition:

$$f(x,y_1) - f(x,y_2) \overset{\text{Appling MVT}}{=} \frac{\partial f}{\partial y}(x,\tilde{y})(y_1 - y_2)$$

Since $\frac{\partial f}{\partial y}$ is continuous on R and thus bounded, we imply

$$|f(x,y_1) - f(x,y_2)| < L \cdot |y_1 - y_2|, \quad \forall (x,y_i) \in R$$

where

$$L = \max\left\{\text{abs}\left(\frac{\partial f}{\partial y}\right)\middle|(x,y) \in R\right\}$$

Theorem 3.4 — **Picard-Lindelof Theorem (existence part).** Suppose $f \in C(R)$ be such that f satisfies the Lipschitz condition, then there exists $a'' \in (0,a]$ such that (**??**) is solvable with $y(x) \in C([\alpha - a'', \alpha + a''])$.

Proof. Consider the complete metric space

$$X = \{y(x) \in C([\alpha - a, \alpha + a]) \mid \beta - b \le y(x) \le \beta + b\},$$

with the mapping $T : X \to X$ defined as

$$(Ty)(x) = \beta + \int_\alpha^x f(t, y(t))\,\mathrm{d}t$$

It suffices to show that T is a contraction, but here we need to estrict a a smaller number as follows:

1. Well-definedness of T: Take $M := \sup\{f(x,y) \mid (x,y) \in R\}$ and construct $a' = \min\{b/M, a\}$. Consider the complete matric space

$$X = \{y(x) \in C([\alpha - a', \alpha + a']) \mid \beta - b \le y(x) \le \beta + b\}$$

which implies that

$$|(Ty)(x) - \beta| \le \left|\int_\alpha^x f(t, y(t))\,\mathrm{d}t\right| \le M|x - \alpha| \le Ma' \le b,$$

i.e., $T(X) \subseteq X$, and therefore $T : X \to X$ is well-defined.

2. Contraction for T: Construct $a'' \in \min\{a', \frac{1}{2L^*}\}$, where L^* is the Lipschitz constant for f. and consider the complete metric space

$$X = \{y(x) \in C([\alpha - a'', \alpha + a'']) \mid \beta - b \leq y(x) \leq \beta + b\}$$

Therefore for $\forall x \in [\alpha - a'', \alpha + a'']$ and the mapping $T : X \to X$,

$$\begin{aligned} |[T(y_1) - T(y_2)](x)| &\leq \left| \int_\alpha^x [f(t, y_2(t)) - f(t, y_1(t))] \, \mathrm{d}t \right| \\ &\leq \int_\alpha^x |f(t, y_2) - f(t, y_1)| \, \mathrm{d}t \leq \int_\alpha^x L^* |y_2(t) - y_1(t)| \, \mathrm{d}t \\ &\leq L^* |x - \alpha| \sup |y_2(t) - y_1(t)| \leq L^* a'' d_\infty(y_2, y_1) \leq \frac{1}{2} d_\infty(y_2, y_1) \end{aligned}$$

Therefore, we imply $d_\infty(Ty_2, Ty_1) \leq \frac{1}{2} d_\infty(y_2, y_1)$, i.e., T is a contraction.

Applying contraction mapping theorem, there exists $y(x) \in X$ such that $Ty = y$, i.e.,

$$y = \beta + \int_\alpha^x f(t, y(t)) \, \mathrm{d}t$$

Thus y is a solution for the IVP (3.11). ■

R On $[\alpha - a'', \alpha + a'']$, we can solve the IVP (3.11) by recursively applying T:

$$\begin{aligned} y_0(x) &= \beta, \qquad \forall x \in [\alpha - a'', \alpha + a''] \\ y_1 &= T(y_0) = \beta + \int_\alpha^x f(t, \beta) \, \mathrm{d}t \\ y_2 &= T(y_1) \\ &\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned}$$

By studying (3.11) on different rectangles, we are able to show the uniqueness of our solution:

Proposition 3.8 Suppose f satisfies the Lipschitz conditon, and y_1, y_2 are two solutions of (3.11), where y_1 is defined on $x \in I_1$, and y_2 is defined on $x \in I_2$. Suppose $I_1 \cap I_2 \neq \emptyset$ and y_1, y_2 share the same initial value condition $y(\alpha) = \beta$. Then $y_1(x) = y_2(x)$ on $I_1 \cap I_2$.

Proof. Suppose $I_1 \cap I_2 = [p, q]$ and let $z := \sup\{x \mid y_1 \equiv y_2 \text{ on } [\alpha, x]\}$. It suffices to show $z = q$. Now suppose on the contrary that $z < q$, and consider the subtraction $|y_1 - y_2|$:

$$y_i = \beta + \int_\alpha^x f(t, y_i)\,\mathrm{d}t \implies |y_1 - y_2| = \left|\int_z^x f(t, y_1) - f(t, y_2)\,\mathrm{d}t\right|.$$

Construct an interval $I^* = [z - \frac{1}{2L^*}, z + \frac{1}{2L^*}] \cap [p, q]$, and let $x_m = \arg\max_{x \in I^*} |y_1(x) - y_2(x)|$, which implies for $\forall x \in I^*$,

$$\begin{aligned}
|y_1(x) - y_2(x)| &= \left|\int_z^x f(t, y_1) - f(t, y_2)\,\mathrm{d}t\right| \\
&\leq \int_z^x |f(t, y_1(t)) - f(t, y_2(t))|\,\mathrm{d}t \\
&\leq L^* \int_z^x |y_1(x) - y_2(x)|\,\mathrm{d}t \\
&\leq L^* |x - z| |y_1(x_m) - y_2(x_m)| \\
&\leq \frac{1}{2} |y_1(x_m) - y_2(x_m)|.
\end{aligned}$$

Taking $x = x_m$, we imply $y_1 \equiv y_2$ for $\forall x \in I^*$, which contradicts the maximality of z. ■

Combining Theorem (3.4) and proposition (3.8), we imply the existence of a unique "maximal" solution for the IVP (3.11), i.e., the unique solution is defined on a maximal interval.

Corollary 3.3 Let $U \subseteq \mathbb{R}^2$ be an open set such that $f(x, y)$ satisfies the Lipschitz condition for any $[a, b] \times [c, d] \subseteq U$, then there exists x_m and x_M in $\overline{\mathbb{R}}$ such that

- The IVP (3.11) admits a solution $y(x)$ for $x \in (x_m, x_M)$, and if y^* is another solution of (3.11) on some interval $I \subseteq (x_m, x_M)$, then $y \equiv y^*$ on I.
- Therefore $y(x)$ is maximally defined; and $y(x)$ is unique.

■ **Example 3.12** Consider the IVP

$$\begin{cases} \dfrac{\mathrm{d}y}{\mathrm{d}x} = x^2 y^{1/5} \\ y(0) = C \end{cases} \implies \frac{\partial f}{\partial y} = \frac{x^2}{5y^{4/5}}.$$

- Taking $U = \mathbb{R} \times (0, \infty)$ implies $y = \left(\frac{4x^3}{15} + c^{4/5}\right)^{5/4}$, defined on $(\sqrt[3]{-15/4c^{4/5}}, \infty)$.
- When $c = 0$, $f(x, y)$ does not satisfy the Lipschitz condition. The uniqueness of solution does not hold.

■

4.2. Monday for MAT3006

Our first quiz will be held on next Wednesday.

Reviewing.

- Picard Lindelof Theorem on ODEs. e.g., consider

$$\begin{cases} \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x}{1-y}, & (x,y) \in G := (-\infty,\infty) \times (-\infty,1) \\ y(0) = 2 \end{cases}$$

 Since $f \in C^1(G)$ satisfies the Lipschitz condition on some closed ball of the point (x_0, y_0), the setting for Picard Lindelof Theorem is satisfied, and the solution is uniquely given by:

$$y = 1 + \sqrt{1-x^2}, \ -1 < x < 1.$$

 Therefore, the maximal interval of existence is given by $(-1,1)$. In order to restrict G to be open to construct a closed ball of (x_0, y_0), we need the initial condition $y(0) \neq 1$.

4.2.1. Generalization into System of ODEs

Formal Setting of System of ODEs. Consider the system of ODEs

$$\begin{cases} y_1'(x) = f_1(x, y_1(x), \dots, y_n(x)) \\ \vdots \\ y_n'(x) = f_n(x, y_1(x), \dots, y_n(x)) \end{cases} \quad \begin{cases} y_1(\alpha) = \beta_1 \\ \vdots \\ y_n(x) = \beta_n \end{cases}$$

It's convenient to denote

$$\boldsymbol{y}(x) = \begin{pmatrix} y_1(x) \\ \vdots \\ y_n(x) \end{pmatrix} \in C(\mathbb{R}, \mathbb{R}^n), \quad \boldsymbol{f}(x, \boldsymbol{y}) = \begin{pmatrix} f_1(x, \boldsymbol{y}) \\ \vdots \\ f_n(x, \boldsymbol{y}) \end{pmatrix}, \quad \boldsymbol{\beta} := \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$$

Here the notation $C(X,Y)$ denotes the set of bounded continuous mapping from X to Y. Therefore we can express the system of ODE as a compact form:

$$\begin{cases} \boldsymbol{y}' = \boldsymbol{f}(x,\boldsymbol{y}) \\ \boldsymbol{y}(\alpha) = \boldsymbol{\beta} \end{cases}$$

Generalization of Picard Lindelof Theorem. Consider the rectangle

$$S = \{(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{R} \times \mathbb{R}^n \mid \alpha - a \le x \le \alpha + a, \beta_i - b_i \le y_i \le \beta_i + b_i, i = 1,\dots,n\}$$

Suppose that

- $\|\boldsymbol{f}(x,\boldsymbol{y})\| \le M, \forall (x,\boldsymbol{y}) \in S$
- $\|\boldsymbol{f}(x,\boldsymbol{y}) - \boldsymbol{f}(x,\boldsymbol{y}')\| \le L \cdot \|\boldsymbol{y} - \boldsymbol{y}'\|$ for $\forall x \in [\alpha - a, \alpha + a]$

Then consider the complete metric space

$$X = \{\boldsymbol{y} \in C([\alpha - a, \alpha + a], \mathbb{R}^n) \mid \beta_i - b_i \le y_i(x) \le \beta_i + b_i\}$$

(Verification of completeness: if Y is complete, then $C(X,Y)$ is complete.) Under this setting, the similar argument gives the Picard-Lindelof for system of ODEs.

Higher Order ODEs. Note that there is a standard way to transform the ODE with higher order derivatives into a system of first order ODEs. Suppose we want to solve the initival value problem

$$\begin{cases} y^{(m)} = f(x, y, y', \dots, y^{(m-1)}) \\ y(\alpha) = \beta_0,\ y'(\alpha) = \beta_1, \dots, y^{(m-1)}(\alpha) = \beta_{m-1} \end{cases}$$

We can define the variables

$$\begin{pmatrix} y_{m-1}(x) \\ \vdots \\ y_1(x) \\ y_0(x) \end{pmatrix} = \begin{pmatrix} y^{(m-1)}(x) \\ \vdots \\ y'(x) \\ y(x) \end{pmatrix}$$

which gives an equivalent system of ODE:

$$\begin{cases} y'_{m-1} = f(x, y_0, \dots, y_{m-1}) \\ y'_{m-2} = y_{m-1} \\ \quad\vdots \\ \quad y'_0 = y_1 \end{cases}, \text{ with } \begin{cases} y_{m-1}(\alpha) = \beta_{m-1} \\ y_{m-2}(\alpha) = \beta_{m-2} \\ \quad\vdots \\ \quad y_0(\alpha) = \beta_0 \end{cases}$$

4.2.2. Stone-Weierstrass Theorem

Under the compact metric space X, the goal is to approximate **any** functions in $C(X)$. For example, under $X = [a, b]$, one can apply Taylor polynomials $p_n(x)$ to approximate differentiable functions:

$$\|f(x) - p_n(x)\|_\infty < \varepsilon, \text{ for large } n.$$

To formally describe the phenomenon for the approximation of **any** functions in $C(X)$, we need to describe the set of approximate functions, which usually obtains a common property:

Definition 4.3 [Algebra] A subset $\mathcal{A} \subseteq C(X)$ (where X is a general space) is an **algebra** if the following holds:

- If $f_1, f_2 \in \mathcal{A}$, then $\alpha f_1 + \beta f_2 \in \mathcal{A}$
- If $f_1, f_2 \in \mathcal{A}$, then $f_1 \cdot f_2 \in \mathcal{A}$

■ **Example 4.1** 1. $\mathcal{A} = C(X)$ is an algebra.

2. $X = [a,b]$, then $\mathcal{A} = P[a,b] = \{\text{All polynomials } p(x)\}$ is an algebra. ■

The goal is to approximate any $f \in C(X)$ by $p \in \mathcal{A}$, i.e., for $\forall f \in C(X)$, there exists $p \in \mathcal{A}$ such that

$$\|f - p\|_\infty < \varepsilon, \ \forall \varepsilon > 0.$$

In other words, we aim to find an algebra $\mathcal{A} \subseteq C(X)$ such that $\overline{\mathcal{A}} = C(X)$, i.e., $\mathcal{A}$ is dense in $C(X)$.

Theorem 4.2 — **Weierstrass Approximation.** $\mathcal{P}[a,b]$ is dense in $C[a,b]$.

Proof. Consider any function $f \in C[0,1]$. By rescaling, assume that $f \in C[0,1]$. By subtracting a linear function $\ell(x)$, assume that $f(0) = f(1) = 0$. Then we extend $f(x)$ into $\mathbb{R}$ by setting $f(x) = 0, \forall x \notin [0,1]$.

- **Step 1: Construction of approximate function**: Consider the **Landaus kernel function**

$$Q_n(x) = \begin{cases} c_n \cdot (1-x^2)^n, & -1 \le x \le 1 \\ 0, & |x| > 1 \end{cases}$$

where c_n is chosen such that $\int Q_n(x)\,\mathrm{d}x = 1$. Then construct the approximation of f by defining

$$p_n(x) := Q_n * f = \int_{-1}^{1} f(x+t) Q_n(t)\,\mathrm{d}t$$

The intuition behind this construction is that as $n \to \infty$, $Q_n(x) \to \delta(x)$, where

$$\delta(x) = \begin{cases} \infty, & x = 0 \\ 0, & x \neq 0 \end{cases} \implies \int_{-1}^{1} f(x+t)\delta(t)\,\mathrm{d}t = f(x).$$

Step 2: Argue that $p_n(x) \in \mathcal{P}[a,b]$: Now it's clear that

$$\begin{aligned}
p_n(x) &= \int_{-1}^{1} f(x+t)Q_n(t)\,\mathrm{d}t && (4.2a)\\
&= \int_{-x}^{1-x} f(x+t)Q_n(t)\,\mathrm{d}t && (4.2b)\\
&= \int_{-1}^{1} f(u)\cdot Q_n(u-x)\,\mathrm{d}u && (4.2c)\\
&= \int_{-1}^{1} f(u)\cdot (1-(u-x)^2)^n\,\mathrm{d}u, && (4.2d)
\end{aligned}$$

where (4.2b) is because that $f = 0$, for $x \notin [0,1]$ and $Q_n = 0$ for $|x| > 1$; (4.2c) is by change of variables; and (4.2d) is by substitution of $Q_n(x)$. Therefore, p_n is still a polynomial of x.

- **Step 3: Construct an upper bound on c_n:** It's clear that

$$\begin{aligned}
c_n^{-1} &= \int_{-1}^{1} (1-x^2)^n\,\mathrm{d}x\\
&= 2\int_{0}^{1} (1-x^2)^n\,\mathrm{d}x\\
&\ge 2\int_{0}^{1} (1-nx^2)\,\mathrm{d}x\\
&\ge 2\int_{0}^{1/\sqrt{n}} (1-nx^2)\,\mathrm{d}x\\
&= 2\left(\frac{1}{\sqrt{n}} - \frac{1}{3\sqrt{n}}\right) > \frac{1}{\sqrt{n}}
\end{aligned}$$

and therefore $c_n < \sqrt{n}$. As a result, for any fixed $\delta \in (0,1)$, we imply

$$Q_n(x) \le \sqrt{n}(1-\delta^2)^n, \qquad \forall x \in [\delta, 1],$$

which implies $Q_n(x) \to 0$ uniformly on $[\delta, 1]$.

- **Step 4: Show that $\|p_n - f\|_\infty \to 0$.** Since f is continuous, for given $\varepsilon > 0$, there exists $\delta \in (0,1)$ such that

$$|f(x) - f(y)| < \varepsilon, \quad \text{when } |x-y| < \delta, x, y \in [0,1].$$

Therefore, for any $x \in [0,1]$, and for sufficiently large n,

$$\begin{aligned}
|p_n(x) - f(x)| &= \left| \int_{-1}^{1} f(x+t) Q_n(t) - \int_{-1}^{1} f(x) Q_n(t)\, \mathrm{d}t \right| && \text{(4.3a)} \\
&\le \int_{-1}^{1} |f(x+t) - f(x)| Q_n(t)\, \mathrm{d}t && \text{(4.3b)} \\
&\le 2M \int_{-1}^{-\delta} Q_n(t)\, \mathrm{d}t + \frac{\varepsilon}{2} \int_{-\delta}^{\delta} Q_n(t)\, \mathrm{d}t + 2M \int_{\delta}^{1} Q_n(t)\, \mathrm{d}t && \text{(4.3c)} \\
&\le 4M\sqrt{n}(1-\delta^2)^n + \frac{\varepsilon}{2} && \text{(4.3d)} \\
&\le \varepsilon && \text{(4.3e)}
\end{aligned}$$

where (4.3c) is by separating the integrand into three parts, and then upper bounding $|f(x+t) - f(x)|$ by $2M := 2\sup_x |f(x)|$ for the integrand $t \in [-1, \delta) \cup (\delta, 1]$, and upper bounding $|f(x+t) - f(x)|$ by $\frac{\varepsilon}{2}$ due to the continuity of f for the integrand $t \in [\delta, \delta]$; (4.3e) is by choosing n sufficiently enough to make $4M\sqrt{n}(1-\delta^2)^n$ sufficiently small.

Therefore $\|p_n - f\|_\infty = \max_{x \in [0,1]} |p_n(x) - f(x)| < \varepsilon$ for large n. The proof is complete.

■

4.5. Wednesday for MAT3006

The quiz will be held on Wednesday.

Reviewing. Let's go through the proof for Weierstrass Theorem quickly.

- Study $Q_n(x) = c_n(1-x^2)^n$ and construct the approximate function

$$p_n(x) = \int_{-1}^{1} Q_n(t) f(x+t)\,\mathrm{d}t$$

- Show that

$$\begin{aligned}
|p_n(x) - f(x)| &\le \int_{-1}^{1} |f(x+t) - f(x)| Q_n(t)\,\mathrm{d}t \\
&= \left(\int_{\delta}^{1} + \int_{\delta}^{-\delta} + \int_{-1}^{-\delta}\right) |f(x+t) - f(x)| Q_n(t)\,\mathrm{d}t \\
&\le 4M\sqrt{n}(1-\delta^2)^n + \int_{\delta}^{-\delta} |f(x+t) - f(x)| Q_n(t)\,\mathrm{d}t \\
&\le 4M\sqrt{n}(1-\delta^2)^n + \varepsilon \cdot \int_{\delta}^{-\delta} Q_n(t)\,\mathrm{d}t \\
&\le 4M\sqrt{n}(1-\delta^2)^n + \varepsilon
\end{aligned}$$

  Therefore, $\|p_n - f\|_\infty \to 0$ as $n \to \infty$.

- Generalization for $\forall g \in C[0,1]$: Recall that we have assumed $f(0) = f(1) = 0$. Now consider the general case, say

$$g(0) = a, \quad g(1) = b.$$

  Consider $f(x) := g(x) - l(x)$, where l is the line segment from $(0,a)$ to $(1,b)$. Then we imply $|f(x) - p_n(x)| < \varepsilon$, i.e.,

$$|g(x) - (p_n(x) + l(x))| < \varepsilon, \quad \forall x.$$

- Generlization for $\forall h \in C[a,b]$: Recall that we have restrict f is continuous on $[0,1]$. For any $h \in C[a,b]$, define $g(x) = h((b-a)x + a)$ for $x \in [0,1]$. Therefore, $g \in C[0,1]$, i.e., $|g(y) - p_n(y)| < \varepsilon, \forall y \in [0,1]$, which implies

$$|h((b-a)y+a) - p_n(y)| < \varepsilon, \quad \forall y \in [0,1]$$

Applying change of variables with $x = (b-a)y + a$, we imply

$$\left|h(x) - p_n\left(\frac{x-a}{b-a}\right)\right| < \varepsilon, \quad \forall x \in [a,b],$$

where $p_n(\cdot)$ is a polynomial function.

4.5.1. Stone-Weierstrass Theorem

The motivation is to generalize the Weierstrass approximation into the space $C(X)$, where (X,d) is a general compact space. Here $C(X) := \{f : X \to \mathbb{R} \text{ is continuous}\}$. Note that

- $C(X)$ has a norm:
$$\|f\|_\infty := \sup\{f(x) \mid x \in X\}$$
This is well-defined, since $f(X) \subseteq \mathbb{R}$ is compact, i.e., closed and bounded.
- $(C(X), d_\infty)$ is complete. The proof follows similarly from the proof that $C[a,b]$ is complete (see Example (2.15)).

R If X is not compact, then the norm $\|\cdot\|_\infty$ is **not** well-defined on $C(X)$, but this norm is still well-defined on the space

$$C_b(X) = \{f : X \to \mathbb{R} \mid f \text{ is continous and bounded}\}.$$

If X is compact, then $C(X) = C_b(X)$.

Definition 4.10 [Separation Property] Let (X,d) be any metric space, and $\mathcal{A} \subseteq C_b(X)$ is an algebra (closed under linear combination and pointwise product), then

1. $\mathcal{A}$ is said to be equipped with the **separation property** if for any $x_1 \neq x_2 \in X$, there exists $f \in \mathcal{A}$ such that $f(x_1) \neq f(x_2)$
2. $\mathcal{A}$ is said to be equipped with the **nonvanishing property** if for any $x \in X$, there exists $f \in \mathcal{A}$ such that $f(x) \neq 0$.

■

■ **Example 4.4** Suppose that $X := S^1 := \{e^{i\theta} \mid \theta \in [0,2\pi]\} \subseteq \mathbb{C} \cong \mathbb{R}^2$, and consider the algebra

$$\mathcal{A} = \langle g \rangle := \operatorname{span}\{1, g, g^2, \dots\}$$

Define $g : S^1 \to \mathbb{R}$ as $g(e^{i\theta}) = \cos\theta$. Note that

1. $\mathcal{A}$ does not satisfy the separation property: take $e^{i\theta}, e^{i(2\pi-\theta)}$
2. However, $\mathcal{A}$ satisfies the nonvanishing property. Consider the special element of $\mathcal{A}$: $f \equiv 1$.

■

Theorem 4.4 — **Stone-Weierstrass Theorem.** Let (X,d) be a compact space, and $\mathcal{A} \subseteq C(X)$ is an algebra. Then $\overline{A} = C(X)$ iff A satisfies both the **nonvanishing** and **separation** property.

Before going through the proof, we establish two lemmas below:

Proposition 4.12 If both f, g belong to the algebra $\mathcal{A}$, then $\max\{f,g\} \in \mathcal{A}$ and $\min\{f,g\} \in \overline{\mathcal{A}}$.

Proof. Since

$$\begin{aligned}
\max\{f,g\} &= \frac{1}{2}(f+g) + \frac{1}{2}|f-g| \\
\min\{f,g\} &= \frac{1}{2}(f+g) - \frac{1}{2}|f-g|,
\end{aligned}$$

it suffices to show $|h| \in \overline{\mathcal{A}}$ given that $h \in \mathcal{A}$.

Let $M = \max\{|h(x)| \mid x \in X\}$. Consider the function (w.r.t. t) $|t| \in C[-M, M]$. By Weierstrass approximation, there exists a polynomial p such that $||t| - p(t)| < \varepsilon$, which implies

$$||h(x)| - p(h(t))| < \varepsilon.$$

Note that $p(h(t))$ is a polynomial of $h(t)$, and therefore an element from the algebra $\mathcal{A}$. Therefore, $|h|$ can be approximated by some element from $\mathcal{A}$, i.e., $|h| \in \overline{\mathcal{A}}$. ■

Proposition 4.13 Let $\mathcal{A} \subseteq C(X)$ be an algebra satisfying the separation property and non-vanishing property. Then for all $x_1 \neq x_2 \in X$, and any $\alpha, \beta \in \mathbb{R}$, there exists $f \in \mathcal{A}$ such that

$$\begin{cases} f(x_1) = \alpha \\ f(x_2) = \beta \end{cases}$$

Proof. By separation property, there exists $h \in \mathcal{A}$ such that $h(x_1) \neq h(x_2)$.

1. We claim that we can construct a new h such that

$$h(x_1) \neq h(x_2), \quad h(x_1) \neq 0, \quad h(x_2) \neq 0 \tag{4.5}$$

   (a) If both $h(x_1), h(x_2) \neq 0$, we have done.
   (b) If not, suppose $h(x_1) = 0$. By non-vanishing property, there eixsts $p \in \mathcal{A}$ such that $p(x_1) \neq 0$. Then some linear transformation of h and p will do the trick. (hint: construct t such that $h \leftarrow h + t \cdot p$ gives the desired result.)

2. Now suppose the requirement (4.5) is met. Consider the function

$$f(x) = ah(x) + bh^2(x) \in \mathcal{A},$$

   where a, b are two parameters to be determined.

   Indeed, it suffices to find a, b such that $f(x_1) = \alpha, f(x_2) = \beta$, or equivalently, solve the linear system

$$f(x_1) = ah(x_1) + bh^2(x_1) = \alpha$$
$$f(x_2) = ah(x_2) + bh^2(x_2) = \beta$$

Since the determinant of the linear system is not equal to 0, a, b can be clearly found.

The proof is complete. ■

Necessity part of the proof. Given that $\mathcal{A}$ has separation and non-vanishing, we aim to show $\overline{\mathcal{A}} = C(X)$.

1. Take any $f \in C(X)$. By proposition (4.13), for any $x, y \in X$, there exists $\phi_{x,y} \in \mathcal{A}$ such that
$$\begin{cases} \phi_{x,y}(x) = f(x) \\ \phi_{x,y}(y) = f(y) \end{cases}.$$
Construct the open set $U_{x,y} = (f - \phi_{x,y})^{-1}((-\varepsilon, \varepsilon))$, i.e.,
$$U_{x,y} = \{t \in X \mid \phi_{x,y}(t) - \varepsilon < f(t) < \phi_{x,y}(t) + \varepsilon\}.$$

2. It's clear that $x, y \in U_{x,y}$. For fixed $y \in X$, the collection $\{U_{x,y}\}_{x \in X}$ forms an open cover of X. By the compactness of X, there exists the finite subcover
$$\{U_{x_1,y}, \ldots, U_{x_N,y}\} \supseteq X.$$
By proposition (4.12), the function $\phi_y := \max\{\phi_{x_1,y}, \ldots, \phi_{x_N,y}\} \in \overline{\mathcal{A}}$. Furthermore, for $\forall x \in X$, we imply there exists some $U_{x_i,y} \ni x$, i.e.,
$$f(x) < \phi_{x_i,y}(x) + \varepsilon \implies f(x) < \phi_y(x) + \varepsilon, \ \forall x \in X.$$

3. Also, consider $V_y = \cap_{i=1}^N U_{x_i,y}$, which is the open set containing y, and $\{V_y\}_{y \in X}$

covers X (why?). Note that for any $x \in V_y$, we imply $x \in U_{x_i,y}, \forall i$, i.e.,

$$\phi_{x_i,y}(x) - \varepsilon < f(x), \quad \forall i \implies \phi_y(x) - \varepsilon < f(x),\ \forall x \in V_y.$$

By the compactness of X again, we take finite subcover $\{V_{y_j}\}_{j=1}^M$ and define

$$\phi(x) := \min\{\phi_{y_1}(x), \dots, \phi_{y_M}(x)\} \in \overline{\mathcal{A}}.$$

Therefore, for any $x \in X$ we imply $x \in V_{y_m}$, i.e.,

$$\phi_{y_m}(x) - \varepsilon < f(x) \implies \phi(x) - \varepsilon < f(x) \tag{4.6}$$

4. Also, from (2) we have obtained $f(x) < \phi_y(x) + \varepsilon$ for $\forall y \in X$. In particular,

$$f(x) < \phi_{y_m}(x) + \varepsilon,\ \forall m = 1, \dots, M \tag{4.7}$$

Combining (4.6) and (4.7), we imply $|\phi(x) - f(x)| < \varepsilon$.

Therefore, we have constructed a function $\phi \in \overline{\mathcal{A}}$ such that $|\phi(x) - f(x)| < \varepsilon$, which implies $f \in \overline{\overline{\mathcal{A}}} = \overline{\mathcal{A}}$. The proof is complete. ■

5.2. Monday for MAT3006

Our first quiz will be held on this Wednesday.

Reviewing. We have shown that the algebra $\mathcal{A} \subseteq C(X)$ with separation, non-vanishing property implies $\overline{\mathcal{A}} = C(X)$.

Now we show that if $\overline{\mathcal{A}} = C(X)$, then the algebra $\mathcal{A}$ has separation, non-vanishing property:

1. Suppose on the contrary that $\mathcal{A}$ is not separating, i.e., there exists $x_1, x_2 \in X$ such that $\phi(x_1) = \phi(x_2)$, $\forall \phi \in \mathcal{A}$.

   By the defintion of closure, it's clear that for given $S \subseteq (X, d)$, $\forall x \in \overline{S}$, there exists a sequence $\{S_n\}$ in S such that $S_n \to x$.

   Construct $f \in C(X)$ defined by $f(x) = d(x, x_1)$. It follows that

   $$f(x_1) = 0, \quad f(x_2) = d(x_2, x_1) := k > 0$$

   Now we claim that $f \notin \overline{\mathcal{A}}$, since otherwise there exists $\{\phi_n\}$ in $\mathcal{A}$ such that $\phi_n \to f$, i.e.,

   $$\phi_n(x_1) \to f(x_1), \quad \phi_n(x_2) \to f(x_2), \quad \phi_n(x_1) = \phi_n(x_2), \forall n,$$

   i.e., $0 = f(x_1) = f(x_2) > 0$.

2. Suppose on the contrary that $\mathcal{A}$ is not non-vanishing, i.e., there exists some $x_0 \in X$ such that $\phi(x_0) = 0, \forall \phi \in \mathcal{A}$. Construct $g \in C(X)$ defined by $g(x) = d(x, x_0) + 1$. Following the similar idea, we can show that there does not exist $\phi_n \in \mathcal{A}$ such that $\phi_n \to g$, i.e., $g \notin \overline{\mathcal{A}}$, which is a contradiction.

Example 5.4 1. Let $X \subseteq \mathbb{R}^n$ be a compact space. Then the polynomial ring

$$\mathbb{R}[x_1, \dots, x_n] = \{\text{Polynomials in } n \text{ variables with coefficients in } \mathbb{R}\}$$

forms a dense set in $C(X)$.

It's clear that the set $\mathbb{R}[x_1, \dots, x_n]$ satisfies the separating and non-vanishing property.

For the special case $n = 1$ and $X = [a,b]$, we get the Weierstrass Approximation Theorem.

2. In particular, when $X = S^1 \subseteq \mathbb{R}^2$, we imply $\mathbb{R}[x,y]$ is dense in $C(S^1)$.

5.2.1. Stone-Weierstrass Theorem in $\mathbb{C}$

Consider the circle $S^1 \subseteq \mathbb{C}$ and the mappings

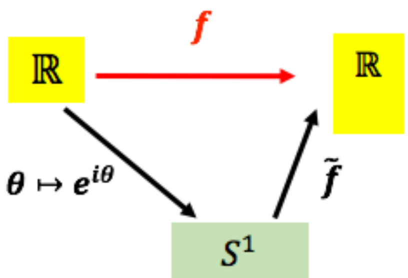
$$\begin{array}{lll} & c: S^1 \to \mathbb{R} & & s: S^1 \to \mathbb{R} \\ \text{with} & e^{i\theta} \to \cos\theta & \text{with} & e^{i\theta} \to \sin\theta \end{array}$$

are both continuous.

The algebra formed by s and c is given by

$$\mathcal{J} := \langle c, s \rangle = \text{span}\{\cos^m\theta \sin^n\theta \mid m,n \in \mathbb{N}\}$$

1. The $\mathcal{J}$ satisfies both separating and non-vanishing property, which implies $\overline{\mathcal{J}} = C(S^1)$.
2. Suppose $f: \mathbb{R} \to \mathbb{R}$ is a continuous, 2π-periodic mapping. It's easy to construct a continuous mapping $\tilde{f}: S^1 \to \mathbb{R}$ such that the diagram below commutes:



Or equivalently, $f(\theta) = \tilde{f}(e^{i\theta})$ for some $\tilde{f} \in C(S^1)$. Since $\overline{\mathcal{J}} = C(S^1)$, we can approximate $\tilde{f} \in C(S^1)$ by $\langle \cos\theta, \sin\theta \rangle$, which implies that the $f(\theta)$ can be approximated by

$$\sum_{m,n\in\mathbb{N}} a_{m,n}\cos^m\theta\sin^n\theta.$$

Since $\text{span}\{\cos^m\theta\sin^n\theta\}_{m,n\in\mathbb{N}} = \text{span}\{\cos(m\theta),\sin(n\theta),1\}_{m,n\in\mathbb{N}}$, we imply $f(\theta)$ can be approximated by

$$\sum_{m,n\in\mathbb{N}} a_m\cos(m\theta)+b_n\sin(n\theta).$$

Or equivalently, for any $\varepsilon>0$, there exists $N>0$ and $a_m, a_n\in\mathbb{R}$ such that

$$\left|f(\theta)-\left(a_0+\sum_{m=1}^{N}a_m\cos(m\theta)+\sum_{n=1}^{N}b_n\sin(n\theta)\right)\right|<\varepsilon,\quad \forall\theta\in[0,2\pi]. \tag{5.1}$$

R The natural question is that do we have the following equation hold:

$$f(\theta)=a_0+\sum_{m=1}^{\infty}a_m\cos(m\theta)+\sum_{n=1}^{\infty}b_n\sin(n\theta) \tag{5.2}$$

It seems that Eq.(5.2) above is equivalent to the expression in (5.4). However, unlike the Taylor expansion, the values of a_m, a_n, M, N may change once we switch the number $\varepsilon>0$.

Therefore, Eq.(5.2) does not hold for most functions, but only for some functions with nice structure.

Fourier Analysis. Given the condition that the Eq.(5.2) holds. How can we get the values of a_m and b_n? The way is to take "inner product" between $f(\theta)$ and trigonometric functions. For example, by taking the inner product with $\cos(k\theta)$ for Eq.(5.2) both sides, we have

$$\begin{aligned}
\int_{-\pi}^{\pi} f(\theta)\cos(k\theta)\,\mathrm{d}\theta &= \frac{a_0}{2}\int_{-\pi}^{\pi}\cos(k\theta)\,\mathrm{d}\theta \\
&\quad+\sum_{m=1}^{\infty}a_m\int_{-\pi}^{\pi}\cos(m\theta)\cos(k\theta)\,\mathrm{d}\theta+\sum_{m=1}^{\infty}b_n\int_{-\pi}^{\pi}\sin(n\theta)\cos(k\theta)\,\mathrm{d}\theta \\
&=\pi\cdot a_k
\end{aligned}$$

Following the same trick, we obtain:

$$\begin{aligned} a_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\cos(k\theta)\,\mathrm{d}\theta \\ b_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\sin(k\theta)\,\mathrm{d}\theta \end{aligned} \tag{5.3}$$

Naturally, we define the fourier expansion for general $f(\theta)$, even though we don't verify whether (5.2) holds or not:

$$g_N(\theta) = \frac{a_0}{2} + \sum_{n=1}^{N} a_m \cos(m\theta) + \sum_{n=1}^{N} b_n \sin(n\theta),$$

where the term a_m and b_n follow the definition in (5.3). The natural question is that whether $g_N(\theta) \to f(\theta)$ as $N \to \infty$?

5.2.2. Baire Category Theorem

Motivation. The set $\mathcal{P}[a,b] \subseteq C[a,b]$ is dense by Weierstrass Approximation. However, it is not "abundant" in $C[a,b]$, just like $\mathbb{Q} \subseteq \mathbb{R}$ is dense in $\mathbb{R}$. (Every $r \in \mathbb{R}$ is a limit of a sequence in $\mathbb{Q}$)

The set $\mathbb{Q}$ is countable yet $\mathbb{R} \setminus \mathbb{Q}$ is uncountable, i.e., there are many more holes in $\mathbb{R} \setminus \mathbb{Q}$.

Definition 5.2 [Nowhere Dense] A subset $S \subseteq (X,d)$ is **nowhere dense** if $\overline{S}$ does not contain any open ball, i.e.,

$$X \setminus \overline{S} \text{ is dense in } X$$

For example, a single point is nowhere dense.

Theorem 5.1 Let $\{E_i\}_{i=1}^{\infty}$ be a collection of nowhere dense sets in a complete metric space (X,d). Then the set

$$\bigcup_{i=1}^{\infty} \overline{E_i}$$

also does not contain any open ball.

Proof. I have no time to review and modify the proof during the lecture. Therefore, we encourage the reader to go through the proof in the note

W,Ni & J. Wang (January, 2019). Lecture Notes for MAT2006. Retrieved from `https://walterbabyrudin.github.io/information/information.html`

Of course, I will also add the proof in this note during this week. ■

5.5. Wednesday for MAT3006

5.5.1. Remarks on Baire Category Theorem

Theorem 5.4 — **Baire Category Theorem.** If (X,d) is complete, and $E_i \subseteq X$ is nowhere dense for $i \in \mathbb{N}$, then

$$\bigcup_{i=1}^{\infty} \overline{E_i}$$

contains no open balls.

Definition 5.4 Let (X,d) be a complete metric space.

1. We say $S \subseteq X$ is **meager** if

$$S = \bigcup_{i=1}^{\infty} E_i, \qquad E_i \text{ is nowhere dense}$$

   In this case we say S is of **first category**.

2. $S' \subseteq X$ is **comeager** if

$$S' = X \setminus S, \qquad \text{where } S \text{ is meager}$$

For example, $\mathbb{Q} = \cup_{x \in \mathbb{Q}} \{x\}$ is megre; $\mathbb{R} \setminus \mathbb{Q}$ is comeager.

R

1. By the Baire Category Theorem, $\cup_{i=1}^{n} \bar{E}_i$ contains no open balls, i.e.,

$$S := \cup_{i=1}^{\infty} E_i \subseteq \cup_{i=1}^{\infty} \bar{E}_i$$

   contains no open balls.
2. S' is comeager implies S' is dense in X: for $\forall x \in X$ and $\forall n \in \mathbb{N}$, $B_{1/n}(x) \cap S'$ is non-empty, since otherwise $X \setminus S'$ contains a open ball, which is a contradiction. Therefore, $x \in \overline{S'}$.

Proposition 5.7 If a set S is meager, it cannot be comeager and vice versa.

Proof. Suppose on contrary that S is meager and comeager, then

$$S = \bigcup_{i=1}^{\infty} E_i, \quad E_i \text{ is nowhere dense}$$
$$X \setminus S = \bigcup_{j=1}^{\infty} F_j, \quad F_j \text{ is nowhere dense}$$

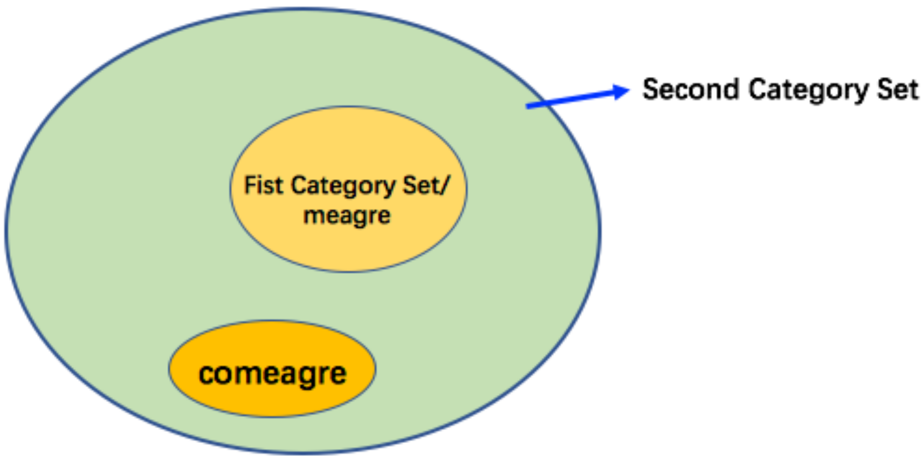
Therefore,

$$X = \bigcup_{i=1}^{\infty} E_i \cup \bigcup_{j=1}^{\infty} F_j$$

is a countable union of nowhere desne sets. By applying Baire Category Theorem, X has no open balls, which is a contradiction. ■

(R) We say $S \subseteq X$ is of **first category** if S is meager. Any subset that is not of first category is of **second category**. Therefore, comeager implies second category.

We illustrate the relationship above in the figure below:



Note that there are subsets that are **neither meager nor co-meager**.

■ **Example 5.6** 1. Here is another proof of $[0,1]$ is un-countable: Suppose on the contrary that $[0,1]$ is countable, then we imply

$$[0,1] = \bigcup_{n\in\mathbb{N}} \{x_n\}, \quad \text{for some } x_n.$$

Applying Baire Category Theorem (since $[0,1]$ is complete), $[0,1] = \cup_{n\in\mathbb{N}}\{x_n\}$ contains no open balls. However, the open ball $(0.5, 0.7) \subseteq [0,1]$, which is a contradiction.

2. The set $X := C[a,b]$ is complete.

   (a) The set of all nowhere differentiable functions is **of 2nd Category** in $C[a,b]$. (Check Theorem (4.1) in MAT2006) Actually, the set of all nowhere differentiable functions is **comeager**. The proof for this statement is omitted.

   (b) Due to the relationship

$$\mathcal{P}[a,b] \subseteq C^{\infty}[a,b] \subseteq \{f : [a,b] \to \mathbb{R} \mid f \text{ is differentiable somewhere}\}$$

   and that the last subset is meager, we imply $\mathcal{P}[a,b]$ and $C^{\infty}[a,b]$ is meager.

■

5.5.2. Compact subsets of $C[a,b]$

Recall that for metrice spaces, the compactness implies closed and bounded, but in general the converse does not hold. We will study extra conditions to make subsets of $C[a,b]$ compact.

Definition 5.5 [(Uniformly) Bounded] The subset S in metric space $(C[a,b], d_\infty)$ is **(uniformly) bounded** if there exists $M > 0$ such that

$$\sup_{f\in S} \|f\|_\infty = M$$

■

In next class, we will show that $K \subseteq C[a,b]$ is compact if and only if K is closed,(uniformly) bounded, and equi-continuous.

6.2. Monday for MAT3006

6.2.1. Compactness in Functional Space

In functional space, previous study have shown that closedness and boundedness is not equivalent to compactness. We need the equi-continuity to rescue the situation:

Definition 6.2 [Equi-continuity] Let $X \subseteq \mathbb{R}^n$. A subset $\mathcal{T} \subseteq C(X)$ is called **equi-continuous** if for any $\varepsilon > 0$, there exists $\delta > 0$ such that whenever $d(x,y) < \delta, x, y \in X$

$$d_\infty(f(x), f(y)) < \varepsilon, \forall f \in \mathcal{T}$$

■ **Example 6.3**

1. Let $\mathcal{T}$ be a collection of Lipschitz continuous functions with the same Lipschitz constant L, i.e., $\forall f \in \mathcal{T}$, $|f(x) - f(y)| < L|x-y|$ for $\forall x, y \in X$. It's clear that $\mathcal{T}$ is equi-continuous.
2. Let $\mathcal{T} \subseteq C[a,b]$ be such that

$$\sup_{x \in [a,b]} |f'(x)| < M, \quad \forall f \in \mathcal{T},$$

then for any $\forall x, y \in [a,b]$, we imply $|f(y) - f(x)| = |f'(\xi)||y-x|$ for some $\xi \in [a,b]$. Therefore,

$$|f(y) - f(x)| < M|y-x|, \quad \forall f \in \mathcal{T},$$

i.e., $\mathcal{T}$ reduces to the space studied in (1) with Lipschitz constant M, thus is equi-continuous.

Theorem 6.3 Let $K \subseteq \mathbb{R}^n$ be a compact set, and $\mathcal{T} \subseteq C(K)$. Then $\mathcal{T}$ is **compact** if and only if $\mathcal{T}$ is **closed**, **uniformly bounded**, and **equicontinuous**.

Proof. To be added. ■

Corollary 6.1 Let $K \subseteq \mathbb{R}^n$ be compact, and $\{f_n\}$ be a sequence of uniformly bounded, equi-continuous functions on K. Then $\{f_n\}$ has the **Bolzano-Weierstrass property**, i.e., it has a convergent subsequence.

Proof. To be added. ■

6.2.2. An Application of Ascoli-Arzela Theorem

The Ascoli-Arzela Theorem has a novel application on the ODE. Consider the IVP problem again:

$$\begin{cases} \dfrac{\mathrm{d}y}{\mathrm{d}x} = f(x,y) \\ y(\alpha) = \beta \end{cases} \tag{6.2}$$

where f is continuous on a rectangle R containing (α,β). Now we show the existence of Picard-Lindelof Theorem without the Lipschitz condition:

Theorem 6.4 — **Cauchy-Peano Theorem.** Consider the problem (6.3). Then there exists a solution of this ODE on some rectangle $R' \subseteq R$.

Proof. To be added. ■

7.2. Monday for MAT3006

Our first mid-term will be held on this Wednesday.

Reviewing. In last lecture, we mainly talk about

- The extended real line
- Definition for limsup and liminf
- For interval I of the form (a,b), $[a,b)$, $(a,b]$ or $[a,b]$, we define

$$m(I) := b - a$$

- We constructed a kind of function to measure the length of a given subset $E \subseteq \mathbb{R}$:

$$m^*(E) = \inf\left\{\sum_{n=1}^{\infty} m(I_n) \middle| E \subseteq \bigcup_{n=1}^{\infty} I_n,\ I_n \text{ are open intervals}\right\}$$

  which is called the **outer measure**.

7.2.1. Remarks on the outer measure

Proposition 7.8
1. $m^*(\phi) = 0$, $m^*(\{x\}) = 0$.
2. $m^*(E + x) = m^*(E)$
3. $m^*(I) = b - a$, where I denotes any interval with endpoints a or b.
4. If $A \subseteq B$, then $m^*(A) \le m^*(B)$
5. $m^*(kE) = |k| m^*(E)$
6. $m^*(\cup_{m=1}^{\infty} E_n) \le \sum_{n=1}^{\infty} m^*(E_n)$ for subsets $E_n \subseteq \mathbb{R}$

R The trick in the proof to show $x \le y$ is by the argument $x \le y + \varepsilon, \forall \varepsilon > 0$.

(1),(2),(5) is clear. (4) is by one-line argument:

Suppose that $B \subseteq \cup_{n=1}^{\infty} I_n$, then $A \subseteq \cup_{n=1}^{\infty} I_n$.

Proof for (3). Consider $m^*([a,b])$ first. The proof for $m^*([a,b]) \leq b-a$ is by explicitly constructing a sequence of open intervals:

$$[a,b] \subseteq (a-\frac{\varepsilon}{2}, b+\frac{\varepsilon}{2}) \cup (a,a) \cup \cdots$$

It follows that

$$\begin{aligned} m^*([a,b]) &\leq m((a-\frac{\varepsilon}{2}, b+\frac{\varepsilon}{2})) + 0 + \cdots + 0 \\ &= (b-a) + \varepsilon, \ \forall \varepsilon > 0 \end{aligned}$$

In particular, $m^*([a,b]) \leq b-a$.

Conversely, the proof for $b-a \leq m^*([a,b])$ is by implicitly constructing a sequence of open interval via the infimum. For all $\varepsilon > 0$, there exists $I_n,\ n \in \mathbb{N}$ such that

$$[a,b] \subseteq \cup_{n=1}^{\infty} I_n, \qquad \sum_{n=1}^{\infty} m(I_n) \leq m^*([a,b]) + \varepsilon.$$

By Heine-Borel Theorem, there exists finite subcover $[a,b] \subseteq \cup_{n=1}^{k} I_n$. Let $I_n = (\alpha_n, \beta_n)$, consider $\alpha := \min\{\alpha_n \mid a \in I_n\}$ and $\beta := \max\{\beta_n \mid b \in I_n\}$. Then we imply

$$[a,b] \subseteq (\alpha, \beta) \subseteq \cup_{n=1}^{k} I_n.$$

It's clear that $\beta - \alpha \leq \sum_{n=1}^{k} m(I_n)$, which follows that

$$b-a \leq \beta-\alpha \leq \sum_{n=1}^{k} m(I_n) \leq \sum_{n=1}^{\infty} m(I_n) \leq m^*([a,b]) + \varepsilon$$

The proof is complete.

The other cases of (3) follows similarly. For example, $m^*((a,b))$ can be lower bounded as:

$$m^*((a,b)) + \varepsilon \geq m^*([a+\frac{\varepsilon}{2}, b-\frac{\varepsilon}{2}]) + \varepsilon = b-a$$

■

Proof for (6). The case for which $m^*(E_n) = \infty$ for some n is trivial, since both sides clearly equal to infinite. Consider the case where $m^*(E_n) < \infty$ only.

By definition, for each E_n we can find $\{I_{n,k}\}_{k=1}^{\infty}$ such that

$$E_n \subseteq \cup_{k=1}^{\infty} I_{n,k}, \quad \sum_{k=1}^{\infty} m(I_{n,k}) \leq m^*(E_n) + \frac{\varepsilon}{2^n}.$$

It follows that

- $\cup_{n=1}^{\infty} \cup_{k=1}^{\infty} I_{n,k}$ is a countable open cover of $\cup_{n=1}^{\infty} E_n$, i.e.,
$$m^*(\cup_{n=1}^{\infty} E_n) \leq \sum_{n,k} m(I_{n,k})$$
- $$\sum_{n,k} m(I_{n,k}) \leq \sum_{n=1}^{\infty} m^*(E_n) + \varepsilon$$

The proof is complete. ■

The natural question is that when does the equality in (6) holds? We will study it in next week.

Definition 7.4 [Null Set] The set $E \subseteq \mathbb{R}$ is a **null set** if $m^*(E) = 0$.

Null sets are the set of points which we can "ignore" when consider the length for sets.

Corollary 7.1
1. If E is null, so is any subset $E' \subseteq E$
2. If E_n is null for all $n \in \mathbb{E}$, so is $\cup_{n=1}^{\infty} E_n$
3. All countable subsets of $\mathbb{R}$ are null.

Proof. (1) follows from (4) in proposition (7.8); (2) follows from (6) in proposition (7.8); (3) follows from (1) and (6) in proposition (7.8). ■

In the remaining of this lecture let's discuss two interesting questions:

1. Are there any uncountable null sets?
2. Both "null" and "meagre" is small. Is null = meagre?

The classic example, cantor set is meagre, null, and uncountable:

■ **Example 7.3** [Cantor Set] Starting from the interval $C_0 = [0,1]$, one delete the open middle third $(1/3,2/3)$ from C_0, leaving two line segments:

$$C_1 = [0,1/3] \cup [2/3,1].$$

Next, the open middle third of each of these remaining segments is deleted, leaving four line segments:

$$C_2 = [0,1/9] \cup [2/9,1/3] \cup [2/3,7/9] \cup [8/9,1].$$

Continuing this process infinitely, and define $C = \cap_{n=1}^{\infty} C_n$.

1. The cantor set C is null, since $C \subseteq C_n$ for all n, i.e.,

$$m^*(C) \le m^*(C_n) = (2/3)^n, \ \forall n \implies m^*(C) = 0.$$

2. The cantor set C is uncountable: every element in C can be expressed uniquely in ternary expression, i.e., only use 0,1,2 as digits. Suppose on the contrary that C is countable, i.e., $C = \{c_n\}_{n\in\mathbb{N}}$. Then construct a new number such that $c \notin \{c_n\}_{n\in\mathbb{N}}$ by diagonal argument.
3. C is nowhere dense, i.e., C is meagre:
   (a) Firstly, C is closed, since intersection of closed sets is closed.
   (b) Suppose on the contrary that $(\alpha,\beta) \subseteq C$ for some open interval (α,β), then $(\alpha,\beta) \subseteq C_n = \sqcup_{k=1}^{2^n}[a_{n,k}, b_{n,k}]$ for all n. Therefore, for any fixed n, $(\alpha,\beta) \subseteq [a_{n,k}, b_{n,k}]$ for some k, which implies

$$\beta - \alpha < b_{n,k} - a_{n,k} = \frac{1}{3^n}, \ \forall n \in \mathbb{N}$$

   Therefore, $\beta - \alpha = 0$, which is a contradiction.

■

R However, the answer for the second question is no. There exists a mergre set S with $m^*(S) = \infty$; and also a null set that is co-meagre. The construction of these examples are left as exercise.

The outer measure m^* is a special measure of the length of a given subset. Now we define the generalized measure of length:

Definition 7.5 [Measure] A meaasure of length for all subsets in $\mathbb{R}$ is a function m satisfying

1. $m(\emptyset) = m(\{x\}) = 0$
2. $m(\{a,b\}) = b - a$
3. $m(A + x) = m(A), \forall x \in \mathbb{R}$
4. If $A \subseteq B$, then $m(A) \le m(B)$
5. $m(kA) = |k| m(A)$
6. If $E_i \cap E_j = \emptyset, \forall i \neq j$, then

$$\sum_{i=1}^{\infty} m(E_i) = m(\cup_{i=1}^{\infty} E_i)$$

Question: m^* satisfies (1) to (5), does m^* satisfies (6) for any subsets? In other words, is outer measure the special case of the definition of measure?

Answer: no.

8.2. Monday for MAT3006

Reviewing. We define the **outer** measure of a subset $E \subseteq \mathbb{R}$ to be

$$m^*(E) = \inf\left\{\sum_{n=1}^{\infty} m(I_n) \middle| E \subseteq \bigcup_{n=1}^{\infty} I_n,\ I_n\text{'s are open intervals}\right\}$$

One Special Property of Outer Measure:

$$m^*(\cup_{n=1}^{\infty} E_n) \le \sum_{n=1}^{\infty} m^*(E_n)$$

8.2.1. Remarks for Outer Measure

We want to make a special hyphothesis become true: If E_n's are disjoint, then

$$m^*(\cup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} m^*(E_n) \tag{8.3}$$

However, (8.3) does not necessary hold for a sequence of disjoint subsets $\{E_n\}$. One counter-example is shown in Example (8.2).

■ **Example 8.2** [Vitali Set] Suppose that $A \subseteq [0,1]$ satisfies the following properties:

- For any $x \in \mathbb{R}$, there exists $q \in \mathbb{Q}$ such that $x + q \in A$.
- If $x, y \in A$ such that $x \neq y$, then $x - y \notin \mathbb{Q}$

In other words, the group $\mathbb{R}$ is partitioned into the cosets of its additive subgroup $\mathbb{Q}$, and the properties above say that A contains exactly one member of each coset of $\mathbb{Q}$. The existence of such A relies on the Axiom of Choice. Moreover, we imply:

- $[0,1] \subseteq \bigcup_{q\in[-1,1]\cap\mathbb{Q}}(A - q)$: since $\forall x \in [0,1]$, there exists $q \in \mathbb{Q}$ s.t. $x + q \in A$, which implies $x \in A - q$. Moreover, we can bound the possible region of q:

$$0 \le x + q \le 1 \implies -x \le q \le 1 - x \implies -1 \le q \le 1$$

- $\bigcup_{q\in[-1,1]\cap\mathbb{Q}}(A-q) \subseteq [-1,2]$: elements in $A-q$ are of the form $x-q, x\in[0,1], q\in[-1,1]$, and therefore $x-q\in[-1,2]$.
- The sets $(A-q)$ are disjoint as q varies, i.e., $(A-q_1)\cap(A-q_2)=\emptyset, \forall q_1\neq q_2\in[-1,1]\cap\mathbb{Q}$: Suppose on the contrary that there exists $y\in(A-q_1)\cap(A-q_2)$, which follows

$$y+q_1, y+q_2\in A, y+q_1\neq y+q_2 \implies (y+q_1)-(y+q_2)=q_1-q_2\notin\mathbb{Q}$$

Suppose on the contrary that (8.3) holds for $\{A-q \mid \forall q\in[-1,1]\cap\mathbb{Q}\}$, then

$$m^*\left(\bigcup_{q\in[-1,1]\cap\mathbb{Q}}(A-q)\right)=\sum_{q\in[-1,1]\cap\mathbb{Q}}m^*(A-q)=\sum_{q\in[-1,1]\cap\mathbb{Q}}m^*(A), \tag{8.4}$$

where the second equality is because that $m^*(A-q)=m^*(A),\forall q$. However,

$$1=m^*([0,1])\le m^*\left(\bigcup_{q\in[-1,1]\cap\mathbb{Q}}(A-q)\right)\le m^*([-1,2])=3 \tag{8.5}$$

From (8.4) we derive the $m^*\left(\bigcup_{q\in[-1,1]\cap\mathbb{Q}}(A-q)\right)$ can either be 0 or ∞, which is a contradiction. ▪

8.2.2. Lebesgue Measurable

Therefore, (8.5) does not hold for some bad subsets of $\mathbb{R}$, which are sets cannot be explicitly described. Let's focus on sets with good behaviour only:

Definition 8.2 [Carathedory Property] A subset $E\subseteq\mathbb{R}$ is **measurable** if

$$m^*(A)=m^*(A\cap E)+m^*(A\setminus E) \tag{8.6}$$

for all subsets $A\subseteq\mathbb{R}$, where E is not assumed to be in A, i.e., $A\setminus E := A\cap E^c$. ▪

R To argue whether (8.6) holds, we essentially suffice to verify the inequality $m^*(A)\ge m^*(A\cap E)+m^*(A\setminus E)$. There are many other equivalent definitions for measurable set $E \subseteq \mathbb{R}$:

1. For any $\varepsilon > 0$, there exists open set $U \supseteq E$ such that

$$m^*(U \setminus E) \le \varepsilon$$

2. Its outer and inner measures are equal:

$$m^*(E) = m_*(E) := \sup\left\{ \sum_{n=1}^{\infty} m(I_n) \middle| \bigcup_{n=1}^{\infty} I_n \subseteq E,\ I_n\text{'s are compact, and disjoint subsets} \right\}$$

Note that the inner measure m_* admits the inequality

$$m_*(\cup_{n=1}^{\infty} E_n) \ge \sum_{n=1}^{\infty} m_*(E_n), \text{ for disjoint } E_n$$

R If $E \subseteq \mathbb{R}$, then for all $B \supseteq E$, we have

$$m^*(B) = m^*(B \cap E) + m^*(B \setminus E) = m^*(E) + m^*(B \setminus E): \tag{8.7}$$

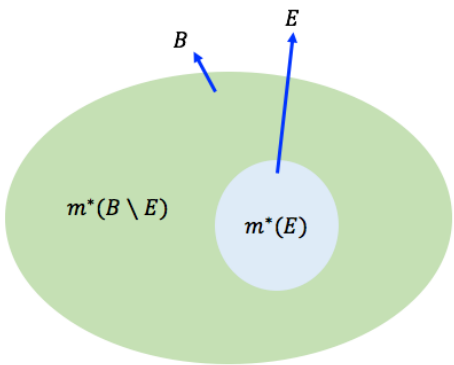


Figure 8.1: Illustration for the useful equality (8.7)

Proposition 8.3 1. If $E \subseteq \mathbb{R}$ is null, then E is measurable

2. If I is any interval, then I is measurable

3. If E is measurable, then $E^c := \mathbb{R} \setminus E$ is measurable

4. If E is measurable, then both $\cup_{i=1}^n E_i$ and $\cap_{i=1}^n E_i$ are measurable

Proof. 1. For any subsets A,

$$\begin{cases} m^*(A \cap E) = 0 \\ m^*(A \cap E^c) \leq m^*(A) \end{cases} \implies m^*(A) \geq m^*(A \cap E) + m^*(A \cap E^c).$$

2. Take $I = [a,b]$. For all $A \subseteq \mathbb{R}$,

   - take $\{I_n\}$ such that $A \subseteq \cup_{n=1}^\infty I_n$ and

$$\sum_{n=1}^{\infty} m^*(I_n) \leq m^*(A) + \varepsilon \tag{8.8}$$

   - Note that the $m^*(A \cap I)$ can be upper bounded:

$$A \cap I \subseteq \cup_{n=1}^\infty (I_n \cap I) \implies m^*(A \cap I) \leq \sum_{n=1}^{\infty} m^*(I_n \cap [a,b])$$

     Similarly, $m^*(A \cap I^c)$ can be upper bounded:

$$A \cap I^c \subseteq \cup_{n=1}^\infty I_n \cap ((-\infty,a) \cup (b,\infty)) = \left(\bigcup_{n=1}^{\infty} I_n \cap (-\infty,a)\right) \cup \left(\bigcup_{n=1}^{\infty} I_n \cap (b,\infty)\right),$$

     i.e.,

$$m^*(A \cap I^c) \leq \sum_{n=1}^{\infty} m^*(I_n \cap (-\infty,a)) + m^*(I_n \cap (b,\infty))$$

   - Therefore,

$$\begin{aligned} m^*(A \cap I) + m^*(A \cap I^c) &\leq \sum_{n=1}^{\infty} m^*(I_n \cap (-\infty,a)) + m^*(I_n \cap [a,b]) + m^*(I_n \cap (b,\infty)) \\ &= \sum_{n=1}^{\infty} m^*(I_n \cap (-\infty,\infty)) = \sum_{n=1}^{\infty} m^*(I_n) \\ &\leq m^*(A) + \varepsilon, \end{aligned}$$

     i.e., $m^*(A \cap I) + m^*(A \cap I^c) \leq m^*(A)$.

3. Part (3) is trivial.

4. Part (4) is by induction on n: suppose that

   - E_i is measurable for $i = 1, \ldots, k+1$
   - $E = \cup_{i=1}^{k} E_i$ is measurable

   By the measurablitiy of E_{k+1},

$$m^*(A \cap E^c) = m^*(A \cap E^c \cap E_{k+1}) + m^*(A \cap E^c \cap E_{k+1}^c) \tag{8.9}$$

   By the measurablitiy of E,

$$\begin{aligned} m^*(A) &\geq m^*(A \cap E) + m^*(A \cap E^c) \\ &\geq [m^*(A \cap E) + m^*(A \cap E^c \cap E_{k+1})] + m^*(A \cap E^c \cap E_{k+1}^c) \end{aligned} \tag{8.10}$$

   It's easy to show

$$E \cup (E^c \cap E_{k+1}) = E \cup E_{k+1},$$

   which implies

$$\begin{aligned} m^*(A \cap (E \cup E_{k+1})) &= m^*(A \cap (E \cup (E^c \cap E_{k+1}))) \\ &= m^*((A \cap E) \cup (A \cap (E^c \cap E_{k+1}))) \\ &\leq m^*(A \cap E) + m^*(A \cap (E^c \cap E_{k+1})) \end{aligned} \tag{8.11}$$

   Substituting (8.11) into (8.10) gives

$$m^*(A) \geq m^*(A \cap (E \cup E_{k+1})) + m^*(A \cap (E \cup E_{k+1})^c),$$

   i.e., $E \cup E_{k+1}$ is measurable as well.

   By the equality

$$\mathbb{R} \setminus \left(\bigcup_{i=1}^{n} E_i \right) = \bigcup_{i=1}^{n} (\mathbb{R} \setminus E_i),$$

   and the result in part (3), one can show $\cap_{i=1}^{n} E_i$ is measurable as well.

■

Proposition 8.4 If E_i is measurable, then $\cup_{i=1}^{\infty} E_i$ is measurable. Moreover, if E_i's are disjoint, then

$$m^*(\cup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} m^*(E_i)$$

R Note that $m^*(A) \neq 0$ for Vitali set A: suppose contrary that $m^*(A) = 0$, i.e., A is null set. Since countably null set is also measurable, together with (8.4), we imply

$$m^*\left(\bigcup_{q\in[-1,1]\cap\mathbb{Q}} (A-q)\right) = 0,$$

which contradicts to (8.5).

Notations.

1. We will write $m(E) = m^*(E)$ for all measurable sets $E \subseteq \mathbb{R}$, and therefore

$$m(\{a,b\}) = m^*(\{a,b\}) = b - a$$

2. The sets E satisfying

$$m^*(A) = m^*(A \cap E) + m^*(A \cap E^c)$$

are called **Lebesgue** measurable in some other textbooks.

8.4. Wednesday for MAT3006

Reviewing.

- All null sets are measurable
- If $E \subseteq \mathbb{R}$ is measurable, then $E^c := R \setminus E$ is measurable.
- E_i is measurable implies $\cup_{i=1}^n E_i$ is measurable.

8.4.1. Remarks on Lebesgue Measurability

Proposition 8.6 If E_i is measurable for $\forall i \in \mathbb{N}$, then so is $\cup_{i=1}^\infty E_i$. Moreover, if further E_i's are pairwise disjoint, then

$$m^*\left(\bigcup_{i=1}^\infty E_i\right) = \sum_{i=1}^\infty m^*(E_i)$$

Proof.
- Consider the case where E_i's are measurable, pairwise disjoint first. For all subsets $A \subseteq \mathbb{R}$, and all $n \in \mathbb{N}$,

$$m^*(A) = m^*(A \cap (\cup_{i=1}^n E_i)) + m^*(A \cap (\cup_{i=1}^n E_i)^c) \tag{8.12a}$$

$$= [m^*(A \cap (\cup_{i=1}^n E_i) \cap E_n) + m^*(A \cap (\cup_{i=1}^n E_i) \cap E_n^c)] + m^*(A \cap (\cup_{i=1}^n E_i)^c) \tag{8.12b}$$

$$= [m^*(A \cap E_n) + m^*(A \cap (\cup_{i=1}^{n-1} E_i)] + m^*(A \cap (\cup_{i=1}^n E_i)^c) \tag{8.12c}$$

where (8.12a) is by the measurability of $\cup_{i=1}^n E_i$; (8.12b) is by the measurability of E_n; (8.12c) is by direct calculation.

Proceeding these trick similarly, we obtain:

$$m^*(A) = [m^*(A \cap E_n) + m^*(A \cap (\cup_{i=1}^n E_i)^c)] + m^*(A \cap (\cup_{i=1}^{n-1} E_i)] \tag{8.13a}$$

$$= \sum_{\ell=1}^n m^*(A \cap E_\ell) + m^*(A \cap (\cup_{i=1}^\ell E_i)^c) \tag{8.13b}$$

$$\geq \sum_{\ell=1}^n m^*(A \cap E_\ell) + m^*(A \cap (\cup_{i=1}^\infty E_i)^c) \tag{8.13c}$$

for any $n \in \mathbb{N}$, where (8.13c) is by lower bounding $(\cup_{i=1}^{\ell} E_i)^c \supseteq (\cup_{i=1}^{\infty} E_i)^c$. Taking $n \to \infty$ in (8.13c), we imply

$$m^*(A) \geq \sum_{\ell=1}^{\infty} m^*(A \cap E_\ell) + m^*(A \cap (\cup_{i=1}^{\infty} E_i)^c) \tag{8.13d}$$

$$\geq m^*(\cup_{i=1}^{\infty}(A \cap E_i)) + m^*(A \cap (\cup_{i=1}^{\infty} E_i)^c) \tag{8.13e}$$

$$= m^*(A \cap (\cup_{i=1}^{\infty} E_i)) + m^*(A \cap (\cup_{i=1}^{\infty} E_i)^c) \tag{8.13f}$$

where (8.13e) is by the countable sub-addictivity of m^*. Therefore, $\cup_{i=1}^{\infty} E_i$ is measurable.

- Moreover, taking $A = \cup_{i=1}^{\infty} E_i$ in (8.13d) gives

$$m^*(\cup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} m^*(E_i) + m^*(\emptyset) = \sum_{i=1}^{\infty} m^*(E_i) + 0.$$

- Now suppose that E_i's are measurable but not necessarily pairwise disjoint. We need to show $\bigcup_{i=1}^{\infty} E_i$ is measurable. The way is to construct the disjoint sequence of sets first:

$$\begin{cases} F_1 = E_1, \\ F_{k+1} = E_k \setminus \left(\cup_{i=1}^{k} E_i\right), \ \forall k > 1 \end{cases} \implies \cup_{i=1}^{\infty} F_i = \cup_{i=1}^{\infty} E_i$$

It's clear that F_i's are pairwise disjoint and measurable, which implies $\cup_{i=1}^{\infty} E_i = \cup_{i=1}^{\infty} F_i$ is measrable. The proof is complete. ■

Notations. We denote $\mathcal{M}$ as the collection of all **(Lebesgue) measurable** subsets of $\mathbb{R}$, and

$$m(E) = m^*(E), \quad \forall E \in \mathcal{M}$$

8.4.2. Measures In Probability Theory

Definition 8.9 [σ-Algebra]

- Let Ω be any set, and $\mathbb{P}(\Omega)$ (**power set**) denotes the collection of all subsets of Ω
- A family of subsets of Ω, denoted as $\mathcal{T}$, is a σ-algebra if it satisfies

    1. $\emptyset, \Omega \in \mathcal{T}$
    2. If $E \in \mathcal{T}$, then $E^c \in \mathcal{T}$
    3. If $E_i \in \mathcal{T}$ for $\forall i \in \mathbb{N}$, then $\cup_{i=1}^{\infty} E_i \in \mathcal{T}$ (and therefore $\cap_{i=1}^{\infty} E_i \in \mathcal{T}$).

Definition 8.10 [Measure] A **measure** on a σ-algerba $(\Omega, \mathcal{T})$ is a function $\mu : \mathcal{T} \to [0, \infty]$ such that

- $\mu(\emptyset) = 0$
- $\mu(\cup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu(E_i)$ whenever E_i's are pairwise disjoint in $\mathcal{T}$.

As a result, $(\Omega, \mathcal{T}, \mu)$ is called a **measurable space**.

Example 8.7 1. Let Ω be any set, $\mathcal{T} = \mathcal{M}$, and $\mu(E) = |E|$ (the number of elements in E). Then $(\Omega, \mathbb{P}(\Omega), \mu)$ is a measure space, and μ is called a counting measure on Ω.

Definition 8.11 [Borel σ-algebra] Let $\boldsymbol{B}$ be a collection of all intervals in $\mathbb{R}$. Then there is a **unique** σ-algebra $\mathcal{B}$ of $\mathbb{R}$, such that

1. $\boldsymbol{B} \subseteq \mathcal{B}$
2. For all σ-algebra $\mathcal{T}$ containing $\boldsymbol{B}$, we have $\mathcal{B} \subseteq \mathcal{T}$

This $\mathcal{B}$ is called a **Borel σ-algebra**

R

1. In particular, $C_i \in \mathcal{B}$ implies $\cup_{i=1}^{\infty} C_i$ and $\cap_{i=1}^{\infty} C_i \in \mathcal{B}$.
2. $\mathcal{B} \subseteq \mathcal{M}$, since $\boldsymbol{B} \subseteq \mathcal{M}$ and $\mathcal{M}$ is a σ-algebra.
3. However, $\mathcal{M}$ and $\mathcal{B}$ are not equal. The element $C \in \mathcal{B}$ is called **Borel measurable subsets**

Definition 8.12 [complete] Let $(\Omega, \mathcal{T}, \mu)$ be a measurable space. Then we say it is **complete** if for any $E \in \mathcal{T}$ with $\mu(E) = 0$, $N \subseteq E$ implies $N \in \mathcal{T}$. (and therefore $\mu(N) = 0$)

Example 8.8

1. $(\mathbb{R}, \mu, m^*)$ is complete.

   Reason: if $m^*(E) = 0$, then $m^*(N) = 0$, $\forall N \subseteq E$
2. $(\mathbb{R}, \mu, m)$ is complete.

   Reason: the same as in (1)
3. However, $(\mathbb{R}, \mathcal{B}, m\mid_{\mathcal{B}})$ is not complete. (left as exercise)

Then we study the difference between $\mathcal{B}$ and $\mathcal{M}$:

Definition 8.13 [Completion] Let $(\Omega, \mathcal{T}, \mu)$ be measurable space. The **completion** of $(\Omega, \mathcal{T}, \mu)$ with respect to μ is the smallest complete σ-algebra containing $\mathcal{T}$, denoted as $\overline{\mathcal{T}}$. More precisely,

$$\overline{\mathcal{T}} = \{G \cup N \mid G \in \mathcal{T}, N \subseteq F \in \mathcal{T}, \text{with } \mu(F) = 0\}$$

e.g., take $G = \emptyset \in \mathcal{T}$. For all $F \in \mathcal{T}$ such that $\mu(F) = 0$, $N \subseteq F$ implies $N \in \overline{\mathcal{T}}$.

R If further define $\overline{\mu} : \overline{\mathcal{T}} \to [0, \infty]$ by

$$\overline{\mu}(G \cup N) = \mu(G),$$

then $(\Omega, \overline{\mathcal{T}}, \overline{\mu})$ is a measurable space.

Theorem 8.3 The completion of $(\mathbb{R}, \mathcal{B}, m\mid_{\mathcal{B}})$ is $(\mathbb{R}, \mathcal{M}, m)$

R Another completion of $(\mathbb{R}, \mu, m)$ is as follows:

Define $\ell(\{a,b\}) = b - a$ for all intervals $\{a,b\} \in \boldsymbol{B}$ Then by Caratheodory extension theorem, we can extend $\ell : \boldsymbol{B} \to [0,\infty]$ to $\ell : \mathcal{B} \to [0,\infty]$.

Complete $\ell : \mathcal{B} \to [0,\infty]$ to $\bar{\ell} : \mathcal{M} \to [0,\infty]$. Then $\bar{\ell} = m$ as in our course.

9.2. Monday for MAT3006

Reviewing.

- The collection of all Lebesgue measurable subsets, denoted by $\mathcal{M}$, is a σ-algebra
- Borel σ-algebra: the smallest σ-algebra containing all the intervals of $\mathbb{R}$:

  Well-definedness. Let $\mathcal{S} = \{\text{all } \sigma\text{-algebras containing all the inervals of } \mathbb{R}\}$. For example, $\mathbb{P}(R) \in \mathcal{S}$. Note that $\forall f_i \in \mathcal{S}$, $\cap_{i\in I}\mathcal{A}_i \in \mathcal{S}$.

  Then define $\mathcal{B} = \cap_{\mathcal{A}\in\mathcal{S}}\mathcal{A}$, which is the smallest σ-algebra containing all intervals.

  ■

  Furthermore, $\mathcal{M} \in \mathcal{S}$. Therefore, $\mathcal{B} \subseteq \mathcal{M}$ but they are not equal. (Check Royden's note on blackboard, there exists a counter-example A such that $A \in \mathcal{M}$ but $A \notin \mathcal{B}$.)

- The set $\mathcal{M}$ has a good property: If $N \in \mathcal{M}$ is null, then all $E \subseteq N$ are null sets, and therefore $E \in \mathcal{M}$.

  The probelm is that it is not necessary the case that $N' \in \mathcal{B}$ implies $m\mid_{\mathcal{B}}(N') = 0$, i.e., $E' \in \mathcal{B}, \forall E' \subseteq N'$ does not necessarily hold.

  (check back to the Roydon's counter-example)

- Therefore, we need the **completion process** of $\mathcal{B}$ to get $\mathcal{M}$.

9.2.1. Measurable Functions

Motivation. The Riemann integration divides the function into a grid of 1 (unit) squares, and then measure the altitude of the function at the center of each square. Therefore, the total "volume" of this function is 1 times the sum of the altitudes.

However, the Lebesgue integraion aims to study the vertical length of the function, and the total volumn of this function is 1 times the sum of the vertical lengths. Riemann integration

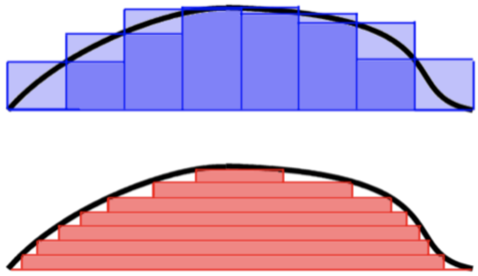


Figure 9.1: Riemann Integration (in blue) and Lebesgue Integration (in red)

Definition 9.1 [Measurable] Let $f : (\mathbb{R}, \mu, m) \to \mathbb{R}$ be a function. We say f is **(Lebesgue) measurable** if $f^{-1}(I) \in \mathcal{M}$ for all intervals $I \subseteq \mathbb{R}$.

Proposition 9.1 If $f : \mathbb{R} \to \mathbb{R}$ is continuous, then f is measurable.

The trick during the proof is to check only intervals of the form (a, ∞) instead of checking all intervals in $\mathbb{R}$.

Proof. 1. By continuity of f, $f^{-1}((a, \infty))$ is open in $\mathbb{R}$.

- Note that any open set U can be expressed as a countable union of open intervals: for given $q \in \mathbb{Q}$, define the set

$$I_q = \bigcup_{I \text{ is an open interval, } q \in I \subseteq U} I,$$

which is a union of non-disjoint open intervals, hence an open interval as well. We claim that $U \subseteq \cup_{q \in \mathbb{Q} \cap U} I_q$:

Consider any $x \in U$. When $x \in \mathbb{Q}$, the result is clear; otherwise there exists an open interval $(x - \varepsilon, x + \varepsilon) \subseteq U$. By the denseness of $\mathbb{Q}$, there exists $q \in \mathbb{Q}$ such that $q \in (x - \varepsilon, x + \varepsilon)$. By definition of I_q, $(x - \varepsilon, x + \varepsilon) \in I_q$. Therefore, $x \in I_q$.

The proof for this statement is complete.

Therefore,

$$f^{-1}((a, \infty)) = \cup_{i=1}^{\infty} U_i \in \mathcal{M}$$

since each open interval $U_i \in \mathcal{M}$

2. For other types of intervals, e.g., $[a, \infty)$, consider

$$\bigcap_{n=1}^{\infty}(a - \frac{1}{n}, \infty) = [a, \infty),$$

which follows that

$$f^{-1}([a, \infty)) = f^{-1}(\bigcap_{n=1}^{\infty}(a - \frac{1}{n}, \infty)) = \cap_{n=1}^{\infty} f^{-1}((a - \frac{1}{n}, \infty)) \in \mathcal{M}$$

The proof for other types of intervals needs similar reformulations of them:

$$f^{-1}((-\infty, a)) = f^{-1}(\mathbb{R} \setminus [a, \infty)) = \mathbb{R} \setminus f^{-1}([a, \infty)) \in \mathcal{M}$$

$$f^{-1}((b, a)) = f^{-1}((-\infty, a)) \cap f^{-1}((b, \infty)) \in \mathcal{M}$$

■

R

1. From the proof above we also find: the function f is measurable if and only if $f^{-1}((a, \infty)) \in \mathcal{M}$, for $\forall a \in \mathbb{R}$.
2. Homework question: the function f is measurable if and only if $f^{-1}(B) \in \mathcal{M}$ for $\forall B \in \mathcal{B}$.

Proposition 9.2 1. Constant functions, and monotone functions are measurable

2. If $A \subseteq \mathbb{R}$ is measurable, then the characterstic function

$$\mathcal{X}_A(x) := \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A \end{cases}$$

is measurable.

3. If f is measurable, h is continuous, then $h \circ f$ is continuous.
4. If f, g are measurable, then so is

$$f + g, \quad fg, \quad \max/\min(f, g), \quad |f|$$

Proof.

- (1) and (2) are easy to show.
- The proof for (3) is simply by applying the formula

$$(h \circ f)^{-1}((a,\infty)) = f^{-1}(h^{-1}(a,\infty)),$$

- The proof for the measurability of $f+g$ is by definition:

$$\begin{aligned}(f+g)^{-1}(a,\infty) &= \{x \mid f+g \in (a,\infty)\} \\ &= \cup_{q\in\mathbb{Z}}(\{x \mid f \in (q,\infty)\} \cap \{x \mid g \in (a-q,\infty)\}) \\ &= \cup_{q\in\mathbb{Z}}(f^{-1}(q,\infty) \cap f^{-1}(a-q,\infty)) \in \mathcal{M}\end{aligned}$$

  The measurability of $fg, |f|, \max/\min(f,g)$ are by the equalities

$$\begin{aligned} fg &= \frac{1}{4}[(f+g)^2 + (f-g)^2] \\ |f| &= h \circ f \quad h(x) = |x| \\ \max/\min(f,g) &= \frac{1}{2}(f+g \pm |f-g|) \end{aligned}$$

∎

(R) If both f,g are measurable, then $g \circ f$ is not necessarily measurable.

Definition 9.2 [Almost Everywhere] Let $f,g : (\mathbb{R},\mu,m) \to \mathbb{R}$. We say $f = g$ almost everywhere (a.e.) if $E := \{x \mid f(x) \neq g(x)\}$ is a null set.

More generally, we say $f(x)$satisfies a condition on (R,μ,m) a.e. if the set

$$\{x \mid f(x) \text{ does not satisfy the condition}\} \text{ is a null set.}$$

For example, the characteristic function $\mathcal{X}_{\mathbb{Q}}(x)$ is equal to zero function a.e.

The measurability ignores the null set.

Proposition 9.3 Suppose that f is measurable, and $g = f$ a.e., then g is measurable.

Proof. Note that

$$g^{-1}((a,\infty)) = \{x \mid g(x) \in (a,\infty), g(x) = f(x)\} \cup \{x \mid g(x) \in (a,\infty), g(x) \neq f(x)\}$$

where $\{x \mid g(x) \in (a,\infty), g(x) \neq f(x)\} \subseteq E$, i.e., belongs to $\mathcal{M}$; and

$$\{x \mid g(x) \in (a,\infty), g(x) = f(x)\} = f^{-1}((a,\infty)) \cap E^c \in \mathcal{M}.$$

∎

(R) During the proof, we have used the fact that $N \subseteq E$ is measurable for all null set E.

Definition 9.3 [Measurable on extended real line] A function $f : \mathbb{R} \to [-\infty,\infty]$ is measurable if $f^{-1}(I) \in \mathcal{M}$ for all intervals $I \in [-\infty,\infty]$.

Following the similar idea of previous examples, it suffices to show that

$$f^{-1}((a,\infty]) \in \mathcal{M}, \quad \forall a \in \mathbb{R}.$$

Or equivalently,

$$f^{-1}(B) \in \mathcal{M}, \forall B \in \mathcal{B}, \text{ and } f^{-1}(\{-\infty\}), f^{-1}(\{\infty\}) \in \mathcal{M}.$$

Example:

$$f(x) = \begin{cases} \tan x & x \neq \dfrac{2n+1}{2}\pi, n \in \mathbb{Z} \\ \infty, & x = \dfrac{2n+1}{2}\pi, n \in \mathbb{Z} \end{cases}$$

is measurable.

9.5. Wednesday for MAT3006

9.5.1. Remarks on Measurable function

Proposition 9.7 Let f_n be a sequence of measurable functions $f_n : \mathbb{R} \to [-\infty, \infty]$. Then the functions

$$\sup_{n\in\mathbb{N}} f_n(x), \quad \inf_{n\in\mathbb{N}} f_n(x), \quad \lim_{n\to\infty} \sup f_n(x), \quad \lim_{n\to\infty} \inf f_n(x)$$

are measurable.

Proof.

- $$\begin{aligned}\left(\sup_{n\in\mathbb{N}} f_n\right)^{-1}((a,\infty]) &= \{x \in \mathbb{R} \mid \sup_n f_n(x) > a\} \\ &= \{x \in \mathbb{R} \mid f_n(x) > a \text{ for some } a\} \\ &= \bigcup_{n\in\mathbb{N}} f_n^{-1}((a,\infty])\end{aligned}$$

  which is measurable due to the measurability of f_n.

- The proof for the measurablitiy of $\inf_n f_n(x), \lim_{n\to\infty} \sup f_n(x), \lim_{n\to\infty} \inf f_n(x)$ is directly by applying the formula

$$\begin{aligned}\inf f_n(x) &= -(\sup(-f_n(x))) \\ \lim_{n\to\infty} \sup f_n(x) &= \lim_{m\to\infty} (\sup_{n\geq m} f_n(x)) = \inf_{m\in\mathbb{N}} (\sup_{n\geq m} f_n(x)) \\ \lim_{n\to\infty} \inf f_n(x) &= -\lim_{n\to\infty} \sup(-f_n(x))\end{aligned}$$

∎

Corollary 9.5 If $\{f_n\}$ is measurable, and $f_n(x)$ converges to $f(x)$ pointwisely a.e., then f is measurable.

Proof. By proposition (9.3), w.l.o.g., $f_n(x)$ conveges to $f(x)$ pointwisely, which follows that

$$f(x) := \lim_{n\to\infty} f_n = \lim_{n\to\infty} \sup f_n(x)$$

i.e., f is measurable due to the measurability of $\lim_{n\to\infty} \sup f_n(x)$. ■

9.5.2. Lebesgue Integration

Definition 9.9 [Simple Function] A function $\phi : \mathbb{R} \to \mathbb{R}$ is **simple** if

- ϕ is measurable and
- $\{\phi(x) \mid x \in \mathbb{R}\}$ takes finitely many values.

More precisely, if the simple function ϕ takes distinct values $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}$ on disjoint non-empty sets $A_1, \ldots, A_k \subseteq \mathbb{R}$, then

$$\phi = \sum_{i=1}^{k} \alpha_i \mathcal{X}_{A_i}$$

Note that A_i's are measurable since $\phi^{-1}(\{\alpha_i\}) = A_i$

R

1. All functions written in the form $\psi = \sum_{i=1}^{\ell} \beta_i \mathcal{X}_{B_i}$, where B_i's are measurable, are simple; All simple functions can be expressed as the form $\psi = \sum_{i=1}^{\ell} \beta_i \mathcal{X}_{B_i}$ (where B_i's are disjoint) uniquely, up to permutation of terms. This is called the canonical form.
2. If ϕ_1, ϕ_2 are simple, then so are

$$\phi_1 + \phi_2, \quad \phi_1 \cdot \phi_2, \quad \alpha \cdot \phi, \max(\phi_1, \phi_2), \quad h \circ \phi.$$

for all function h.

Definition 9.10 [Lebesgue integral for Simple Function] Given a simple function with the canonical form $\phi := \sum_{i=1}^{k} \alpha_i \mathcal{X}_{\mathcal{A}_i}$,

- The Lebesgue integral for ϕ (over $\mathbb{R}$) is

$$\int \phi \, \mathrm{d}m = \sum_{i=1}^{k} \alpha_i m(A_i),$$

- The Lebesgue integral for ϕ over a measurable set E is

$$\int_E \phi \, \mathrm{d}m = \int \phi \cdot \mathcal{X}_E \, \mathrm{d}m = \sum_{i=1}^{k} \alpha_i m(A_i \cap E)$$

Proposition 9.8 For any simple function $\phi = \sum_{i=1}^{\ell} \beta_i \mathcal{X}_{B_i}$, where B_i's are not necessarily disjoint, we still have

$$\begin{aligned}\int \phi \, \mathrm{d}m &= \sum_{i=1}^{\ell} \beta_i m(B_i), \\ \int (\phi + \psi) \, \mathrm{d}m &= \int \phi \, \mathrm{d}m + \int \psi \, \mathrm{d}m, \quad \text{where } \psi \text{ is another simple function,} \\ \int \phi \, \mathrm{d}m &\le \int \psi \, \mathrm{d}m, \qquad \text{provided that } \phi \le \psi.\end{aligned}$$

Proof. It suffices to show the first equality. w.l.o.g., suppose $\phi = \beta_1 \mathcal{X}_{B_1} + \beta_2 \mathcal{X}_{B_2}$, which can be reformulated as the canonical form:

$$\phi = (\beta_1 + \beta_2)\mathcal{X}_{B_1 \cap B_2} + \beta_1 \mathcal{X}_{B_1 \cap B_2^c} + \beta_2 \mathcal{X}_{B_1^c \cap B_2}$$

Then we can take the Lebesgue integration for ϕ:

$$\int \phi \, \mathrm{d}m = (\beta_1 + \beta_2) m(B_1 \cap B_2) + \beta_1 m(B_1 \cap B_2^c) + \beta_2 m(B_1^c \cap B_2),$$

which is equal to $\beta_1 m(B_1) + \beta_2 m(B_2)$ due to the caratheodory property (definition (8.2))

■

Definition 9.11 [Lebesgue integral for Measurable Function] Let f be a measurable function $f : \mathbb{R} \to [0, \infty]$. Then the Lebesgue integral of f is given by:

$$\int f \, dm = \sup \left\{ \int \phi \, dm \middle| 0 \le \phi \le f, \phi \text{ is simple} \right\} \tag{9.3}$$

We say f is integrable if $\int f \, dm < \infty$.

R

- It's not appropriate if we try to define the Lebesgue integral by

$$\int f \, dm = \inf \left\{ \int \phi \, dm \middle| 0 \le f \le \phi, \phi \text{ is simple} \right\} \tag{9.4}$$

The problem is due to the function $f(x) = \frac{1}{\sqrt{x}}$ on $(0, 1)$. Note that the function values can be arbitrarily large.

Since a simple function takes only finitely many values, every simple function that is bounded below by f has to be infinite on a set of non-zero measure.

Therefore, the integral using your suggested infimum definition would be ∞, whereas the usual Lebesgue integral would have a finite value.

- Also, one can try to define $\int f \, dm$ for non-measurable function f. The problem is that

$$\int (f + g) \, dm \ne \int f \, dm + \int g \, dm \text{ in general}$$

We will see the detailed reason later.

Proposition 9.9
- The formula (9.3) and (9.4) matches with each other for any simple functions $\phi \ge 0$.
- For $\alpha \ge 0$,

$$\int \alpha f \, dm = \alpha \int f \, dm$$

- If $0 \le f \le g$, then

$$\int f \, \mathrm{d}m \le \int g \, \mathrm{d}m$$

Proof. omitted. ■

Proposition 9.10 — **Markov Inequality.** Suppose that $f : \mathbb{R} \to [0, \infty]$ is measurable, then

$$m(f^{-1}[\lambda, \infty]) \le \frac{1}{\lambda} \int f \, \mathrm{d}m$$

Corollary 9.6 If $f : \mathbb{R} \to [0, \infty]$ is integrable, then $m(f^{-1}\{\infty\}) = 0$, i.e., f is finite a.e.

Proof.

$$m(f^{-1}\{\infty\}) \le m(f^{-1}[\lambda, \infty]) \le \frac{1}{\lambda} \int f \, \mathrm{d}m, \ \forall \lambda \ge 0.$$

Since $\int f \, \mathrm{d}m$ is finite, we imply $\frac{1}{\lambda} \int f \, \mathrm{d}m$ can be arbitrarily small, i.e., $m(f^{-1}\{\infty\}) = 0$. ■

10.2. Monday for MAT3006

10.2.1. Remarks on Markov Inequality

Proposition 10.1 — **Markov Inequality.** Suppose that $f:\mathbb{R}\to[0,\infty]$ is measurable, then

$$m(f^{-1}[\lambda,\infty])\le\frac{1}{\lambda}\int f\,\mathrm{d}m,\ \forall\lambda>0$$

Proof. Define the function

$$g:=\lambda\mathcal{X}_{f^{-1}([\lambda,\infty])},$$

it follows that $g\le f$ globally. Applying proposition (9.9), we imply

$$\int g\,\mathrm{d}m\le\int f\,\mathrm{d}m\implies\lambda m(f^{-1}[\lambda,\infty])\le\int f\,\mathrm{d}m.$$

■

Corollary 10.1 If $f:\mathbb{R}\to[0,\infty]$ is integrable, and $\int f\,\mathrm{d}m=0$, then $f=0$ a.e.

Proof. Consider that for any $\lambda>0$,

$$0\le m(f^{-1}[\lambda,\infty])\le\frac{1}{\lambda}\int f\,\mathrm{d}m=0.$$

Therefore, $m(\{x\mid f(x)\neq0\})=m(f^{-1}(0,\infty])=0$. ■

10.2.2. Properties of Lebesgue Integration

In this lecture, we will show several lemmas, which is very useful during the proof of monotone convergence theorem.

Proposition 10.2 If $f:\mathbb{R}\to[0,\infty]$ is such that $f=0$ a.e., then $\int f\,\mathrm{d}m=0$.

Proof. Any simple function $\psi \leq f$ must be 0 almost everywhere:

$$\phi = \sum_i \alpha_i \mathcal{X}_{A_i}, \alpha_i > 0, \cup_i A_i \text{ is null.}$$

Direct computation of the Lebesgue integral for this simple function ψ gives

$$\int f \, \mathrm{d}m = \sum_i \alpha_i m(A_i) = 0,$$

where the last equality is because that for each i, the set A_i is null. ■

R Given a non-negative integrable function f on a measurable set E, the integral $\int_E f \, \mathrm{d}m = 0$ if and only if $f = 0$ a.e. on E.

Proposition 10.3 If A, B are measurable, disjoint sets, then

$$\int_{A \cup B} f \, \mathrm{d}m = \int_A f \, \mathrm{d}m + \int_B f \, \mathrm{d}m$$

Proof. The key is to apply $f \cdot \mathcal{X}_{A \cup B} = f \cdot \mathcal{X}_A + f \cdot \mathcal{X}_B$ and

$$\int_E f \, \mathrm{d}m = \int f \cdot \mathcal{X}_E \, \mathrm{d}m, \text{ for any measurable } E.$$

■

Proposition 10.4 If $f : \mathbb{R} \to [0, \infty]$ is measurable, then there exists an increasing sequence of simple functions $\{\phi_n\}$ such that $\phi_n(x) \to f(x)$ pointwise.

Proof. For each $n \in \mathbb{N}$, we divide the interval $[0, 2^n] \subseteq [0, \infty]$ into 2^{2n} subintervals of width 2^{-n}:

$$I_{k,n} = (k2^{-n}, (k+1)2^{-n}], \quad k = 0, 1, \dots, 2^{2n} - 1.$$

Let $J_n = (2^n, \infty]$ be the remaining part of the range of f, and define

$$E_{k,n} = f^{-1}(I_{k,n}), \quad F_n = f^{-1}(J_n).$$

Then the sequence of simple functions are given by:

$$\phi_n = \sum_{k=0}^{2^n-1} k \cdot 2^{-n} \mathcal{X}_{E_{k,n}} + 2^n \mathcal{X}_{F_n}.$$

■

Proposition 10.5 — Fatou's Lemma. Let $\{F_n\}$ be a sequence of non-negative measurable functions, then

$$\liminf_{n\to\infty} \int f_n \,\mathrm{d}m \geq \int \left(\liminf_{n\to\infty} f_n\right) \mathrm{d}m$$

R The inequality in the Fatou's lemma could be strict, e.g., consider $f_n(x) = (n+1)x^n$ on $[0,1]$.

10.5. Wednesday for MAT3006

Proposition 10.12 — **Fatou's Lemma.** Suppose $\{f_n\}$ is a sequence of measurable, non-negative functions.

$$\liminf_{n\to\infty}\int f_n \,\mathrm{d}m \ge \int \liminf_{n\to\infty}(f_n)\,\mathrm{d}m$$

Proof. Define $g_n(x) := \inf_{k\ge n} f_k(x)$ and

$$f(x) = \liminf_{n\to\infty} f_n(x) = \lim_{n\to\infty}\left(\inf_{k\ge n} f_k(x)\right) := \lim_{n\to\infty} g_n(x)$$

To study the integral $\int f\,\mathrm{d}m$, we will only focus on $f(x)$ on $E\subseteq\mathbb{R}$, where $f(x)>0, \forall x\in E$.

It suffices to show that $\int_E \phi\,\mathrm{d}m \le \lim_{n\to\infty}\inf\int_E f_n\,\mathrm{d}m$ for all simple ϕ satisfying $0\le\phi(x)\le f(x), \forall x\in E$. (Then taking supremum both sides leads to the desired result.)

1. Construct the simple function ϕ' on E such that

$$\phi'(x) = \begin{cases} \phi(x)-\varepsilon, & \text{if } \phi(x)>0 \\ 0, & \text{if } \phi(x)=0 \end{cases}$$

   in which we pick ε small enough such that $\phi(x)-\varepsilon\ge 0$.

   As a result, $\phi' < f, \forall x\in E$ (why?).

2. Note that $g_n(x)$ is monotone increasing with n, and therefore convergent to $f(x)$. Consider $A_n := \{x\in E \mid \phi'(x)\le g_n(x)\}$, which follows that

   (a) $A_n\subseteq A_{n+1}$

   (b) $\cup_{n=1}^{\infty} A_n = E$ (We do need ϕ' is strictly less than f to obtain this condition).

   Therefore, for any $k\ge n$,

$$\int_{A_n}\phi'\,\mathrm{d}m \le \int_{A_n} g_n\,\mathrm{d}m \le \int_{A_n} f_k\,\mathrm{d}m,$$

   which implies $\int_{A_n}\phi'\,\mathrm{d}m \le \int_E f_k\,\mathrm{d}m$ since $f_k\mathcal{X}_{A_n}\le f_k\mathcal{X}_E$. Or equivalently,

$$\int_{A_n}\phi'\,\mathrm{d}m \le \inf_{k\ge n}\int_E f_k\,\mathrm{d}m \tag{10.2}$$

3. Taking limits $n \to \infty$ both sides for (10.2):

   - For LHS, suppose that $\phi' = \sum_i \alpha_i \mathcal{X}_{c_i}$, then $\int_{A_n} \phi' \, \mathrm{d}m = \sum_i \alpha_i m(c_i \cap A_n)$, which follows that

$$\lim_{n\to\infty} \int_{A_n} \phi' \, \mathrm{d}m = \sum_i \alpha_i \lim_{n\to\infty} m(c_i \cap A_n) = \sum_i \alpha_i m(c_i) = \int_E \phi' \, \mathrm{d}m$$

   - The limit of RHS equals $\lim_{n\to\infty} \inf \int_E f_n \, \mathrm{d}m$, and therefore

$$\int_E \phi' \, \mathrm{d}m \le \lim_{n\to\infty} \inf \int_E f_n \, \mathrm{d}m$$

   Note that the goal is to show $\int_E \phi \, \mathrm{d}m \le \lim_{n\to\infty} \inf \int_E f_n \, \mathrm{d}m$, and therefore we need to evaluate ϕ' in terms of ϕ.

4. (a) Consider the case where $m(\phi^{-1}(0,\infty)) = P < \infty$, then

$$\int_E \phi' \, \mathrm{d}m = \int_E \phi \, \mathrm{d}m - \varepsilon \cdot P \le \lim_{n\to\infty} \inf \int_E f_n \, \mathrm{d}m,$$

   for all small $\varepsilon > 0$. Then the desired result holds.

   (b) Consider the case where $m(\phi^{-1}(0,\infty)) = \infty$, and we write the canonical form $\phi = \sum \alpha_i \mathcal{X}_{c_i}$ with $\alpha_i > 0$. Define $C = \cup_i c_i$ such that $m(c) = \infty$.
   Construct the simple function $\phi' = a\mathcal{X}_C$, where $a := \frac{1}{2}\min\{\alpha_i\}$, which implies

   - $\phi' \le \phi$
   - $\int_E \phi' \, \mathrm{d}m = am(c) = \infty$, which follows that $\int_E \phi \, \mathrm{d}m = \infty$.

   Our goal is to show $\lim_{n\to\infty} \inf \int_E f_n \, \mathrm{d}m = \infty$.
   Consider $B_n = \{x \in E \mid g_n(x) > a\}$, then $\cup B_n = E, B_n \subseteq B_{n+1}$.
   Observe the inequality

$$\int_{C\cap B_n} a \, \mathrm{d}m \le \int_{B_n} a \, \mathrm{d}m \le \int_{B_n} g_n \, \mathrm{d}m \le \inf_{k\ge n} \int_E f_n \, \mathrm{d}m$$

   Taking $n \to \infty$ both sides. For LHS, by definition of B_n, the limit equals $\int_C a \, \mathrm{d}m = \int \phi' \, \mathrm{d}m = \infty$; and the limit of RHS equals to $\lim_{n\to\infty} \inf \int_E f_n \, \mathrm{d}m$, i.e.,

$$\liminf_{n\to\infty} \int_E f_n \,\mathrm{d}m = \infty$$

∎

Theorem 10.3 — **Monotone Convergence Theorem I.** Let $\{f_n\}$ be a sequence of non-negative measurable functions, with

- $f_n(x)$ being monotone increasing
- $f_n(x) \to f(x)$ pointwisely

Then we have

$$\lim_{n\to\infty} \int f_n \,\mathrm{d}m = \int \left(\lim_{n\to\infty} f_n\right) \mathrm{d}m := \int f \,\mathrm{d}m$$

Proof.
- On the one hand, for all $n \in \mathbb{N}$, we have

$$f_n \le f \implies \int f_n \,\mathrm{d}m \le \int f \,\mathrm{d}m \implies \limsup_{n\to\infty} \int f_n \,\mathrm{d}m \le \int f \,\mathrm{d}m$$

- On the other hand, applying the Fatou's lemma,

$$\int f \,\mathrm{d}m := \int \left(\liminf_{n\to\infty} f_n\right) \mathrm{d}m \le \liminf_{n\to\infty} \int f_n \,\mathrm{d}m$$

  Togehter with the previous inequality, we imply

$$\limsup_{n\to\infty} \int f_n \,\mathrm{d}m \le \int f \,\mathrm{d}m \le \liminf_{n\to\infty} \int f_n$$

Therefore, all inequalities above are equalities, and the limit exists since limsup and liminf coincides. Moreover,

$$\lim_{n\to\infty} \int f_n \,\mathrm{d}m = \int f \,\mathrm{d}m.$$

∎

From MCT I, the Lebesgue integral $\int f \,\mathrm{d}m$ can be computed as follows:

- Construct simple functions $\phi_n \le \phi_{n+1}$ with $\phi_n \to f$

- Evaluate $\int \phi_n \, dm$ and then $\int f \, dm = \lim_{n\to\infty} \int \phi_n \, dm$

10.5.1. Consequences of MCT

Proposition 10.13 The Lebesgue integral is finitely addictive for measurable non-negative functions. In other words, suppose f, g are measurable and nonnegative, then

$$\int f \, dm + \int g \, dm = \int (f+g) \, dm$$

Proof. Suppose we have simple increasing functions $\{\phi_n\}$ and $\{\psi_n\}$ such that $\phi_n \to f$ and $\psi_n \to f$. Then

$$\int (f+g) \, dm = \lim_{n\to\infty} \int (\phi_n + \psi_n) \, dm \tag{10.3a}$$

$$= \lim_{n\to\infty} \int \phi_n \, dm + \lim_{n\to\infty} \int \psi_n \, dm \tag{10.3b}$$

$$= \int f \, dm + \int g \, dm \tag{10.3c}$$

where (10.3a) and (10.3c) is by applying MCT I; and (10.3b) is by definition of simple function. ∎

Corollary 10.2 The Lebesgue integral is linear defined for measurable, nonnegative functions. In other words, suppose f, g are measurable and nonnegative, then

$$\int (af + bg) \, dm = a \int f \, dm + b \int g \, dm,$$

for any $a, b \ge 0$.

Proposition 10.14 The Lebesgue integral for non-negative continuous function on a bounded closed interval coincides with the Riemann integral. In other words, let f be a non-negative continuous function on $[a,b]$. then

$$\int_{[a,b]} f\,\mathrm{d}m = \int_a^b f(x)\,\mathrm{d}x.$$

We will extend this result into all proper Riemann integrable functions on $[a,b]$ soon.

Proof. Let ϕ_n be the simple function giving the Riemann lower sum of $f(x)$ with 2^n equal subintervals:

$$\phi_n(x) = \sum_{k=1}^{2^n} \left(\min_{y\in \bar{I}_k} f(y)\right) \mathcal{X}_{I_k}, \text{ where } I_k = [a+(b-a)\frac{k-1}{2^n}, a+(b-a)\frac{k}{2^n}]$$

- $\phi_n(x) \geq 0$ is monotone increasing (that's the reason we should divde intervals into 2^n pieces instead of n pieces)

- $\phi_n(x) \to f(x)$ pointwisely: for any $x \in [a,b]$ and $\varepsilon > 0$, by (uniform) continuity of f, there exists $\delta > 0$ such that

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

  Therefore, for sufficiently large n, we imply for any $x \in I_{k,n}$, $|I_{k,n}| < \delta$. As a result,

$$\left|\min_{y\in I_{k,n}} f(y) - f(x)\right| < \varepsilon.$$

Therefore,

$$\begin{aligned}
\int_{[a,b]} f\,\mathrm{d}m &= \lim_{n\to\infty} \int \phi_n\,\mathrm{d}m \\
&= \lim_{n\to\infty} \left[\text{Riemann lower integral of } \int_a^b f(x)\,\mathrm{d}x\right] \\
&= \int_a^b f(x)\,\mathrm{d}x
\end{aligned}$$

■

Example 10.3 The Lebesgue integral gives us an alternative way to compute improper integrals. Suppose that we want to compute the integral

$$\int_0^1 (1-x)^{-1/2}\,\mathrm{d}x.$$

1. The old method is that we know the integral

$$\int_0^{1-1/n} (1-x)^{-1/2}\,\mathrm{d}x \text{ exists for any } n.$$

Then we extend the definition of Riemann integration by taking limit of n:

$$\int_0^1 f(x)\,\mathrm{d}x = \lim_{n\to\infty}\int_0^{1-1/n}(1-x)^{-1/2} = \lim_{n\to\infty} 2-2\sqrt{\frac{1}{n}} = 2$$

2. The Lebesgue integration does not require us to extend the definition. Consider

$$f_n(x) = (1-x)^{-1/2}\mathcal{X}_{[0,1-1/n]}$$

Then

- $f_n(x) \to f(x)$ on $[0,1)$
- $f_n(x)$ is monotone increasing

Therefore, by applying MCT I,

$$\int_{[0,1)} (1-x)^{-1/2}\,\mathrm{d}x = \lim_{n\to\infty}\int f_n\,\mathrm{d}m = \lim_{n\to\infty}\int_{[0,1-1/n]}(1-x)^{-1/2}\,\mathrm{d}x.$$

By Proposition (10.14), $\int_{[0,1-1/n]}(1-x)^{-1/2}\,\mathrm{d}x = \int_0^{1-1/n}(1-x)^{-1/2}\,\mathrm{d}x$ for all n. Therefore, we conclude that the Lebesgue integral is equal to the (improper) Riemann integral in this case.

11.2. Monday for MAT3006

Reviewing. Compute the integration

$$\int_{[0,1)} (1-x)^{-1/2}\,\mathrm{d}x$$

Solution. 1. Construct $g_n(x) = (1-x)^{-1/2}\mathcal{X}_{[0,1-1/n]}$, then g_n is monotone increasing and $g_n(x) \to (1-x)^{-1/2}\mathcal{X}_{[0,1)}$ pointwisely.

2. By applying MCT I and proposition (10.14),

$$\int_{[0,1)} (1-x)^{-1/2}\,\mathrm{d}x = \lim_{n\to\infty}\int g_n\,\mathrm{d}m = 2.$$

■

Question: How to understand $\int_{[0,1]}(1-x)^{-1/2}\,\mathrm{d}x$?

Answer:

$$(1-x)^{-1/2}\mathcal{X}_{[0,1]} = (1-x)^{-1/2}\mathcal{X}_{[0,1)} + \infty\cdot\mathcal{X}_{\{1\}}$$

which follows that

$$\begin{aligned}\int (1-x)^{-1/2}\mathcal{X}_{[0,1]}\,\mathrm{d}m &= \int (1-x)^{-1/2}\mathcal{X}_{[0,1)}\,\mathrm{d}m + \int \infty\cdot\mathcal{X}_{\{1\}}\,\mathrm{d}m && (11.1a)\\ &= \int_{[0,1)}(1-x)^{-1/2}\,\mathrm{d}x + 0 && (11.1b)\end{aligned}$$

where (11.1b) is because that $\infty\cdot\mathcal{X}_{\{1\}} = \infty\cdot 0 = 0$.

11.2.1. Consequences of MCT I

Proposition 11.4 If f, g are measurable non-negative functions, and $f = g$ a.e., then

$$\int f\,\mathrm{d}m = \int g\,\mathrm{d}m$$

Proof. Let $U = \{x \in \mathbb{R} \mid f(x) = g(x)\}$, then

$$f = f \cdot \mathcal{X}_U + f \cdot \mathcal{X}_{U^c}$$

where U^c is null. As a result,

$$\begin{aligned}
\int f \, \mathrm{d}m &= \int f\mathcal{X}_U + \int f\mathcal{X}_{U^c} && (11.2a)\\
&= \int g\mathcal{X}_U + 0 && (11.2b)\\
&= \int g\mathcal{X}_U + \int g\mathcal{X}_{U^c} && (11.2c)\\
&= \int g \, \mathrm{d}m && (11.2d)
\end{aligned}$$

where (11.2b) is because that $f \cdot \mathcal{X}_{U^c} = 0$ a.e., and $f \cdot \mathcal{X}_U = g \cdot \mathcal{X}_U$; (11.2c) is becasue that $g \cdot \mathcal{X}_{U^c} = 0$ a.e. ∎

Proposition 11.5 — Slight Generalization of MCT I. Suppose that $f_n(x)$ are nonnegative measurable functions such that

1. f_n is monotone increasing a.e.
2. $f_n(x) \to f(x)$ a.e.

then

$$\lim_{n\to\infty} \int f_n \, \mathrm{d}m = \int f \, \mathrm{d}m$$

Proof. Construct the set $V_n = \{x \mid f_n(x) \le f_{n+1}(x)\}$ and $V = \bigcap_{n=1}^{\infty} V_n$. Since $f_n(x)$ is monotone increasing a.e., we imply $m(V_n^c) = 0$, and $m(V^c) \le \sum_{n=1}^{\infty} m(V_n^c) = 0$.

1. Construct $\tilde{f}_n(x)$ as follows:

$$\tilde{f}_n(x) = \begin{cases} f_n(x), \text{ if } x \in V \\ 0, \text{ if } x \in V^c \end{cases}$$

   As a result,

   - $\tilde{f}_n$ is monotone increasing
   - Define a function $g : \mathbb{R} \to [0, \infty]$ such that $\lim_{n\to\infty} \tilde{f}_n(x) = g_n(x)$.

Apply the MCT I gives

$$\lim_{n\to\infty}\int \tilde{f}_n \,\mathrm{d}m = \int g \,\mathrm{d}m \tag{11.3a}$$

2. Note that $\{x \mid \tilde{f}_n(x) \neq f_n(x)\} \subseteq V^c$, where V^c is null. Therefore, $f_n = f$ a.e., which implies

$$\int \tilde{f}_n \,\mathrm{d}m = \int f_n \,\mathrm{d}m \tag{11.3b}$$

3. Consider $V' = \{x \mid \lim_{n\to\infty} f_n(x) = f(x)\}$, and $(V')^c$ is null by hyphothesis. For any $x \in V \cap V'$, we imply

$$f(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \tilde{f}_n(x).$$

Since $(V \cap V')^c$ is null, we imply $\tilde{f}_n(x) \to f$ a.e. Note that $\tilde{f}_n(x) \to g$, we imply $g = f$ a.e., which follows that

$$\int g \,\mathrm{d}m = \int f \,\mathrm{d}m \tag{11.3c}$$

Combining (11.3a) to (11.3c), we conclude that

$$\lim_{n\to\infty}\int f_n \,\mathrm{d}m = \lim_{n\to\infty}\int \tilde{f}_n \,\mathrm{d}m = \int g \,\mathrm{d}m = \int f \,\mathrm{d}m$$

■

Proposition 11.6 Let $\{f_k\}$ be non-negative measurable and

$$f := \sum_{k=1}^{\infty} f_k,$$

then

$$\int f \,\mathrm{d}m = \sum_{k=1}^{\infty}\int f_k \,\mathrm{d}m$$

Proof. Firstly, $\int f \,\mathrm{d}m$ is well-defined since $f = \lim_{n\to\infty}\sum_{k=1}^{n} f_k$ is measurable.

Secondly, take $g_n = \sum_{k=1}^{n} f_k$, which implies g_n is monotone increasing and $g_n \to f$. Apply MCT I gives the desired result. ■

■ **Example 11.3** Consider

$$(1-x)^{-1/2} = \sum_{n=0}^{\infty} \frac{(2n)!}{4^n (n!)^2} x^n, \quad x \in [0,1)$$

Take $f_k = \frac{(2k)!}{4^k (k!)^2} x^k$. Applying proposition (11.6) gives

$$\int_{[0,1)} (1-x)^{-1/2} \, \mathrm{d}x = \sum_{n=0}^{\infty} \int_0^1 \frac{(2n)!}{4^n \cdot (n!)^2} x^n \, \mathrm{d}x$$

Or equivalently,

$$2 = \sum_{n=0}^{\infty} \frac{(2n)!}{4^n (n!)(n+1)!}$$

■

11.2.2. MCT II

We now extend our study to all measurable functions instead of non-negativity.

Definition 11.4 [Lebesgue integrable] Let f be a measurable function, then let

$$f^+(x) = \begin{cases} f(x), & \text{if } f(x) > 0 \\ 0, & \text{if } f(x) \le 0 \end{cases} = f(x)\mathcal{X}_{f^{-1}((0,\infty])}$$

and

$$f^-(x) = \begin{cases} -f(x), & \text{if } f(x) \le 0 \\ 0, & \text{if } f(x) > 0 \end{cases} = -f(x)\mathcal{X}_{f^{-1}([-\infty,0])}$$

As a result, f^+ and f^- are both measurable.

Note that

- $f(x) = f^+(x) - f^-(x)$
- $|f|(x) = f^+(x) + f^-(x)$

Now we define the Lebesgue integral of f as

$$\int f \,\mathrm{d}m = \int f^+ \,\mathrm{d}m - \int f^- \,\mathrm{d}m$$

We say f is **Lebesgue integrable** if both f^+ and f^- are integrable, i.e., $\int f^\pm \,\mathrm{d}m < \infty$

Proposition 11.7

1. If f is measurable, then f is integrable if and only if $|f|$ is integrable
2. If f is measurable, and $|f| \leq g$ with g integrable, then f is also integrable

Proof.

1. If f is integrable, then $\int f^+ \,\mathrm{d}m, \int f^- \,\mathrm{d}m < \infty$. As a result,

$$\int |f| \,\mathrm{d}m = \int (f^+ + f^-) \,\mathrm{d}m = \int f^+ \,\mathrm{d}m + \int f^- \,\mathrm{d}m < \infty.$$

For the reverse direction, if $|f|$ is integrable, then

$$\int |f| = \int f^+ + \int f^-$$

therefore $\int f^\pm < \infty$, and hence f is interable.

2. Since $0 \leq |f| \leq g$, by proposition (9.9), $\int |f| \,\mathrm{d}m \leq \int g \,\mathrm{d}m < \infty$.

Therefore, $\int |f| \,\mathrm{d}m < \infty$, and hence $|f|$ is integrable, which implies f is integrable.

■

R If $|f| \leq g$, and $\int |f| \,\mathrm{d}m = \infty$, then by proposition (9.9), we imply $\int g \,\mathrm{d}m = \infty$.

11.5. Wednesday for MAT3006

Proposition 11.12 — **Linearity.** If f, g are both integrable, then $f+g$ and αf are integrable with

$$\int (f+g)\,\mathrm{d}m = \int f\,\mathrm{d}m + \int g\,\mathrm{d}m$$
$$\int \alpha f\,\mathrm{d}m = \alpha \int f\,\mathrm{d}m \quad \alpha \in \mathbb{R}$$

Proof. 1. Construct

$$(f+g)^+ - (f+g)^- = f+g = (f^+ - f^-) + (g^+ - g^-) \implies (f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+$$

Since both sides for the equality above is non-negative, we do the Lebesgue integral both sides:

$$\int ((f+g)^+ + f^- + g^-)\,\mathrm{d}m = \int ((f+g)^- + f^+ + g^+)\,\mathrm{d}m.$$

Due to the linearity of Lebesgue integral for non-negative functions,

$$\int (f+g)^+\,\mathrm{d}m + \int f^-\,\mathrm{d}m + \int g^-\,\mathrm{d}m = \int (f+g)^-\,\mathrm{d}m + \int f^+\,\mathrm{d}m + \int g^+\,\mathrm{d}m$$

i.e.,

$$\int (f+g)\,\mathrm{d}m = \int f\,\mathrm{d}m + \int g\,\mathrm{d}m$$

2. Assume $\alpha < 0$. Then

$$\begin{aligned}
\int (\alpha f)\,\mathrm{d}m &:= \int (\alpha f)^+\,\mathrm{d}m - \int (\alpha f)^-\,\mathrm{d}m \\
&= \int (-\alpha) f^-\,\mathrm{d}m - \int (-\alpha) f^+\,\mathrm{d}m \\
&= (-\alpha)\int f^-\,\mathrm{d}m - (-\alpha)\int f^+\,\mathrm{d}m \\
&= \alpha\left(\int f^+\,\mathrm{d}m - \int f^-\,\mathrm{d}m\right) \\
&= \alpha \int f\,\mathrm{d}m
\end{aligned}$$

The proof for the case $\alpha \geq 0$ follows similarly.

∎

11.5.1. Properties of Lebesgue Integrable Functions

Corollary 11.2 Suppose that f, g are integrable, then

1. If $f \leq g$, then $\int f \, \mathrm{d}m \leq \int g \, \mathrm{d}m$
2. If $f = g$ a.e., then $\int f \, \mathrm{d}m = \int g \, \mathrm{d}m$

Proof. 1. Since $g - f \geq 0$, $\int (g-f) \, \mathrm{d}m \geq \int 0 \, \mathrm{d}m = 0$. By linearity, $\int g \, \mathrm{d}m - \int f \, \mathrm{d}m \geq 0$, i.e.,

$$\int g \, \mathrm{d}m \geq \int f \, \mathrm{d}m.$$

2. The proof follows similarly as in proposition (11.4). In detail, let $U = \{x \mid f(x) = g(x)\}$, then $m(U) = 0$. It follows that

$$\int f \mathcal{X}_{U^C} \, dm = \int f^+ \mathcal{X}_{U^c} \, dm + \int f^- \mathcal{X}_{U^c} \, dm = 0$$

Similarly, $\int g \mathcal{X}_{U^c} \, dm = 0$. Therefore,

$$\begin{aligned}
\int f \, \mathrm{d}m &= \int f \mathcal{X}_U \, \mathrm{d}m + \int f \mathcal{X}_{U^c} \, \mathrm{d}m \\
&= \int g \mathcal{X}_U \, \mathrm{d}m \\
&= \int g \mathcal{X}_U \, \mathrm{d}m + \int g \mathcal{X}_{U^c} \, \mathrm{d}m \\
&= \int g \, \mathrm{d}m
\end{aligned}$$

∎

R

1. Consider the set of integrable functions, say $\mathcal{T} = \{f : \mathbb{R} \to [-\infty, \infty], \text{ integrable}\}$, which is a vector space if we define $0_{\mathcal{T}} :=$ zero function.

We can define a "norm" on $f \in \mathcal{T}$ by

$$\|f\| = \int |f| \, \mathrm{d}m$$

then $\|\alpha f\| = |\alpha| \|f\|$ and $\|f + g\| \le \|f\| + \|g\|$.

Unfortunately, we should keep in mind that $\mathcal{T}$ is not a normed space, since there exists $f \neq 0_{\mathcal{T}}$ such that $\|f\| = 0$, e.g., $f = \mathcal{X}_{\mathbb{Q}}$.

2. To remedy this, define the equivalence relation on $\mathcal{T}$: $f \sim g$ if $f = g$ a.e. The equivalence classes of $\mathcal{T}$ under $\sim$ are of the form $[f] := \{g : g \sim f\}$. Denote the collection of equivalence classes as $L^1(\mathbb{R}) := \mathcal{T}/\sim$.

   (a) It's clear that $L^1(\mathbb{R})$ has a vector space structure

   $$[f] + [g] = [f + g]$$
   $$\alpha[f] = [\alpha f]$$

   (b) The space $L^1(\mathbb{R})$ can be viewed as a quotient space defined in linear algebra. Consider a vector subspace $\mathcal{N}$ of $\mathcal{T}$ defined by

   $$\mathcal{N} := \{g \in \mathcal{T} \mid g = 0 \text{ a.e.}\}$$

   then $\mathcal{T}/\sim = \mathcal{T}/\mathcal{N}$.

   (c) We define a norm on $L^1(\mathbb{R})$ by $\|[f]\| = \int |f| \, \mathrm{d}m$, which is truly a norm:

   $$\begin{aligned} \|\alpha[f]\| &= |\alpha| \|[f]\| \\ \|[f] + [g]\| &\le \|[f]\| + \|[g]\| \\ \|[f]\| = 0 &\iff \int |f| \, \mathrm{d}m = 0 \iff f = 0 \text{ a.e.} \iff [f] = 0_{L^1(\mathbb{R})} \end{aligned}$$

   Similarly, we can study $L^2(\mathbb{R}), \dots, L^p(\mathbb{R})$, e.g., for $L^2(\mathbb{R}) = \{f : \mathbb{R} \to [-\infty, \infty] \mid \int |f|^2 \, \mathrm{d}m < \infty\}/\mathcal{N}$, define the norm

   $$\|f\|_2 = \left(\int |f|^2 \, \mathrm{d}m\right)^{1/2}$$

■ **Example 11.10** There exist some improper Riemann integrable functions that are not Lebesgue integrable: Consider $f = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} \mathcal{X}_{[k,k+1)}$, then the improper Riemann integral gives

$$\int_0^{\infty} f(x)\,\mathrm{d}x = \log(2) = 1 - \frac{1}{2} + \frac{1}{3} - \cdots$$

However, f is not Lebesgue integrable. Suppose on the contrary that it is , then $|f|$ is integrable:

$$|f| = \sum_{k=0}^{\infty} \frac{1}{k+1} \mathcal{X}_{[k,k+1)}$$

However,

$$\begin{aligned}\int |f|\,\mathrm{d}m &= \lim_{n\to\infty} \sum_{k=0}^{n} \int \left(\frac{1}{k+1}\mathcal{X}_{[k,k+1)}\right)\mathrm{d}m \\ &= \sum_{k=0}^{\infty} \frac{1}{k+1} = \infty\end{aligned}$$

We will also show that all the proper Riemann integrable functions are Lebesgue integrable (and the integrals have the same value) ■

Theorem 11.3 — **MCT II.** Let $\{f_n\}$ be a sequence of integrable functions such that

1. $f_n \le f_{n+1}$ a.e.
2. $\sup_n \int f_n\,\mathrm{d}m < \infty$

Then f_n converges to an integrable function f a.e., and

$$\int f\,\mathrm{d}m = \lim_{n\to\infty} \int f_n\,\mathrm{d}m$$

Proof. Re-define f_n by changing its values on a null set such that

1. $f_n(x) \in \mathbb{R}$, for any $x \in \mathbb{R}$
2. $f_n(x) \le f_{n+1}(x)$ for any $n \in \mathbb{R}, x \in \mathbb{R}$

Let $f(x) = \lim_{n\to\infty} f(x)$. Consider the sequence of functions $\{f_n - f_1\}_{n\in\mathbb{N}}$, then

1. $f_n - f_1 \ge 0$

2. $f_n - f_1$ is monotone increasing, integrable
3. $f_n - f_1 \to f - f_1$

Applying MCT I gives

$$\int (f - f_1)\,\mathrm{d}m = \lim_{n\to\infty} \int (f_n - f_1)\,\mathrm{d}m$$

Adding $\int f_1\,\mathrm{d}m$ and applying the linearity of integrals, we obtain

$$\int (f - f_1)\,\mathrm{d}m + \int f_1\,\mathrm{d}m = \lim_{n\to\infty} \int (f_n - f_1)\,\mathrm{d}m + \int f_1\,\mathrm{d}m = \lim_{n\to\infty} \int f_n\,\mathrm{d}m$$

Here $\lim_{n\to\infty} \int f_n\,\mathrm{d}m$ exists as $\lim_{n\to\infty} \int f_n\,\mathrm{d}m = \sup_n \int f_n\,\mathrm{d}m < \infty$; and $\int (f - f_1)\,\mathrm{d}m + \int f_1\,\mathrm{d}m$ is integrable since it equals $\lim_{n\to\infty} \int f_n\,\mathrm{d}m < \infty$.

Therefore,

$$\text{LHS} = \int f\,\mathrm{d}m = \text{RHS} = \lim_{n\to\infty} \int f_n\,\mathrm{d}m.$$

The proof is complete. ∎

12.2. Monday for MAT3006

12.2.1. Remarks on MCT

■ **Example 12.2** The MCT can help us to compute the integral

$$\lim_{n\to\infty}\int_{[0,n\pi]}\cos\left(\frac{x}{2n}\right)xe^{-x^2}\,\mathrm{d}x$$

Construct $f_n(x)=\cos\left(\frac{x}{2n}\right)xe^{-x^2}\mathcal{X}_{[0,n\pi]}$.

- Since $\cos(x/2n)<\cos(x/2(n+1))$ for any $x\in[0,n\pi]$, we imply f_n is monotone increasing with n
- $f_n(x)$ is integrable for all n.
- f_n converges pointwise to $xe^{-x^2}\mathcal{X}_{[0,\infty)}$

Therefore, MCT I applies and

$$\lim_{n\to\infty}\int_{[0,n\pi]}\cos\left(\frac{x}{2n}\right)xe^{-x^2}\,\mathrm{d}x=\int\left(\lim_{n\to\infty}f_n\right)\mathrm{d}m$$

with

$$\lim_{n\to\infty}f_n=xe^{-x^2}\mathcal{X}_{[0,\infty)}.$$

Moreover,

$$\begin{aligned}\int\left(\lim_{n\to\infty}f_n\right)\mathrm{d}m&=\lim_{m\to\infty}\int_{[0,m]}xe^{-x^2}\,\mathrm{d}x &\text{(12.1a)}\\ &=\int_0^\infty xe^{-x^2}\,\mathrm{d}x &\text{(12.1b)}\\ &=\frac{1}{2} &\text{(12.1c)}\end{aligned}$$

where (12.1a) is by applying MCT I with $g_m(x)=xe^{-x^2}\mathcal{X}_{[0,m]}$ and proposition (10.14) to compute a Lebesgue integral by evaluating a proper Riemann integral. ■

Then we discuss the Lebesgue integral for series:

Corollary 12.3 [Lebesgue Series Theorem] Let $\{f_n\}$ be a series of measurable functions such that

$$\sum_{n=1}^{\infty}\int |f_n|\,\mathrm{d}m < \infty,$$

then $\sum_{n=1}^{k} f_n$ converges to an integrable function $f = \sum_{n=1}^{\infty} f_n$ a.e., with

$$\int f\,\mathrm{d}m = \sum_{n=1}^{\infty}\int f_n\,\mathrm{d}m$$

Proof.

- For each f_n, consider

$$f_n = f_n^+ - f_n^-, \text{ where } f_n^+, f_n^- \text{ are nonnegative.}$$

By proposition (11.6),

$$\int \sum_{n=1}^{\infty} f_n^+\,\mathrm{d}m = \sum_{n=1}^{\infty}\int f_n^+\,\mathrm{d}m \le \sum_{n=1}^{\infty}\int |f_n|\,\mathrm{d}m < \infty.$$

Therefore, $f^+ := \sum_{n=1}^{\infty} f_n^+ = \lim_{k\to\infty}\sum_{n=1}^{k} f_n^+$ is integrable. The same follows by replacing f^+ with f^-. By corollary (9.6), $f^+(x), f^-(x) < \infty, \forall x \in U$, where U^c is null.

- Therefore, construct

$$f(x) = \begin{cases} f^+(x) - f^-(x), & x \in U \\ 0, & x \in U^c \end{cases}$$

Moreover, for $x \in U$,

$$\begin{aligned} f(x) &= \left(\lim_{k\to\infty}\sum_{n=1}^{k} f_n^+(x)\right) - \left(\lim_{k\to\infty}\sum_{n=1}^{k} f_n^-(x)\right) \\ &= \lim_{k\to\infty}\left(\sum_{n=1}^{k} f_n^+(x) - \sum_{n=1}^{k} f_n^-(x)\right) \\ &= \lim_{k\to\infty}\left[\sum_{n=1}^{k}(f_n^+(x) - f_n^-(x))\right] \\ &= \sum_{n=1}^{\infty} f_n(x) \end{aligned}$$

where the first equality is because that both terms are finite.

- It follows that

$$\int f\,\mathrm{d}m = \int f^+\,\mathrm{d}m - \int f^-\,\mathrm{d}m \tag{12.2a}$$
$$= \int \sum_{n=1}^{\infty} f_n^+\,\mathrm{d}m - \int \sum_{n=1}^{\infty} f_n^-\,\mathrm{d}m \tag{12.2b}$$
$$= \left(\sum_{n=1}^{\infty} \int f_n^+\,\mathrm{d}m\right) - \left(\sum_{n=1}^{\infty} \int f_n^-\,\mathrm{d}m\right) \tag{12.2c}$$
$$= \sum_{n=1}^{\infty} \left(\int f_n^+\,\mathrm{d}m - \int f_n^-\,\mathrm{d}m\right) \tag{12.2d}$$
$$= \sum_{n=1}^{\infty} \int f_n\,\mathrm{d}m \tag{12.2e}$$

where (12.2a),(12.2d) is because that summation/subtraction between series holds when these series are finite; (12.2c) is by proposition (11.6); (12.2e) is by definition of f_n.

■

Example 12.3 Compute the integral

$$\int_{(0,1]} e^{-x} x^{\alpha-1}\,\mathrm{d}x,\ \alpha > 0.$$

- Construct $f_n(x) = (-1)^n \frac{x^{\alpha+n-1}}{n!} \mathcal{X}_{(0,1]}, n \geq 0$, and

$$\sum_{n=0}^{N} f_n(x) \to e^{-x} x^{\alpha-1},\ \text{pointwisely}, x \in (0,1].$$

By applying MCT I,

$$\int |f_n|\,\mathrm{d}m = \frac{1}{(\alpha+n)n!}$$

Therefore,

$$\sum_{n=0}^{\infty} \int |f_n|\,\mathrm{d}m = \sum_{n=0}^{\infty} \frac{1}{(\alpha+n)n!} < \infty$$

- Applying the Lebesgue Series Theorem,

$$\int_{(0,1]} e^{-x}x^{\alpha-1}\,\mathrm{d}x = \int_{(0,1]} (\sum_{n=0}^{\infty} f_n)\,\mathrm{d}m = \sum_{n=0}^{\infty}\int f_n\,\mathrm{d}m = \sum_{n=0}^{\infty}\frac{(-1)^n}{(\alpha+n)n!}$$

R It's essential to have $\sum\int|f|\,\mathrm{d}m<\infty$ rather than $\sum\int f_n\,\mathrm{d}m<\infty$ in the Lebesgue Series Theorem. For example, let

$$f_n = \frac{(-1)^{n+1}}{(n+1)}\mathcal{X}_{[n,n+1)} \implies \sum_{n=1}^{\infty}\int f_n\,\mathrm{d}m = \log(2) < \infty$$

However, $f := \sum f_n$ is not integrable.

12.2.2. Dominated Convergence Theorem

Theorem 12.2 Let $\{f_n\}$ be a sequence of measruable functions such that $|f_n|\le g$ a.e., and g is integrable. Suppose that $\lim_{n\to\infty} f_n(x) = f(x)$ a.e., then

1. f is integrable,
2. $$\int f\,\mathrm{d}m = \lim_{n\to\infty}\int f_n\,\mathrm{d}m$$

Proof.
- Observe that

$$|f_n|\le g \implies \lim_{n\to\infty}|f_n|\le g \implies |f|\le g$$

By comparison test, g is integrable implies $|f|$ is integrable, and further f is integrable.

- Consider the sequence of non-negative functions $\{g-f_n\}_{n\in\mathbb{N}}$ and $\{g+f_n\}_{n\in\mathbb{N}}$.

By Fatou's Lemma,

$$\begin{aligned}
\liminf_{n\to\infty}\int (g-f_n)\,\mathrm{d}m &\geq \int \liminf_{n\to\infty}(g-f_n)\,\mathrm{d}m\\
&= \int (g-f)\,\mathrm{d}m\\
&= \int g\,\mathrm{d}m - \int f\,\mathrm{d}m
\end{aligned}$$

which follows that

$$\int g\,\mathrm{d}m - \limsup_{n\to\infty}\int f_n\,\mathrm{d}m \geq \int g\,\mathrm{d}m - \int f\,\mathrm{d}m$$

i.e.,

$$\int f\,\mathrm{d}m \geq \limsup_{n\to\infty}\int f_n\,\mathrm{d}m$$

- Similarly,

$$\liminf_{n\to\infty}(g+f_n)\,\mathrm{d}m \geq \int \liminf_{n\to\infty}(g+f_n)\,\mathrm{d}m = \int g\,\mathrm{d}m + \int f\,\mathrm{d}m$$

which implies

$$\liminf_{n\to\infty}\int f_n\,\mathrm{d}m \geq \int f\,\mathrm{d}m$$

As a result,

$$\limsup_{n\to\infty}\int f_n\,\mathrm{d}m \leq \int f\,\mathrm{d}m \leq \liminf_{n\to\infty}\int f_n\,\mathrm{d}m,$$

which implies

$$\int f\,\mathrm{d}m = \lim_{n}\int f_n\,\mathrm{d}m$$

■

Corollary 12.4 [Bounded Convergence Theorem] Suppose that $E \in \mathcal{M}$ be such that $m(E) < \infty$. If

- $|f_n(x)| \leq K < \infty$ for any $x \in E, n \in \mathbb{N}$

- $f_n \to f$ a.e. in E,

then f is integrable in E with

$$\int_E f \, \mathrm{d}m = \lim_{n\to\infty} \int f_n \, \mathrm{d}m$$

Proof. Take $g = K\mathcal{X}_E$ in DCT. ■

Proposition 12.2 Every Riemann integrable function f on $[a,b]$ is Lebesgue integrable, without the condition that f is continuous a.e.

Proof. Since f is Riemann integrable, we imply f is bounded. We construct the Riemann lower abd upper functions with 2^n equal intervals, denoted as $\{\phi_n\}$ and $\{\psi_n\}$, which follows that

- ϕ_n is monotone increasing; ψ_n is monotone decreasing;
- $\phi_n \le f \le \psi_n$, and

$$\lim_{n\to\infty} \int_{[a,b]} \phi_n = \int_a^b f(x) \, \mathrm{d}x = \lim_{n\to\infty} \int_{[a,b]} \psi_n.$$

Construct $g = \sup_n \phi_n$ and $h = \inf_n \psi_n$. Now we can apply the bounded convergence theorem:

- ϕ_n is bounded on $[a,b]$
- $\phi_n \to g$ on $[a,b]$

which implies g is Lebesgue integrable on $[a,b]$, with

$$\int_{[a,b]} g \, \mathrm{d}m = \lim_{n\to\infty} \int_{[a,b]} \phi_n \, \mathrm{d}m = \int_a^b f(x) \, \mathrm{d}x.$$

Similarly, h is Lebesgue integrable, with

$$\int_{[a,b]} h \, \mathrm{d}m = \lim_{n\to\infty} \int_{[a,b]} \psi_n \, \mathrm{d}m = \int_a^b f(x) \, \mathrm{d}x.$$

Moreover, $g \leq f \leq h$, and

$$\int_{[a,b]} (h-g)\,\mathrm{d}m = \int_{[a,b]} h\,\mathrm{d}m - \int_{[a,b]} g\,\mathrm{d}m = \int_a^b f(x)\,\mathrm{d}x - \int_a^b f(x)\,\mathrm{d}x = 0,$$

which implies $h = g$ a.e., and further $f = g$ a.e., which implies

$$\int_{[a,b]} f\,\mathrm{d}m = \int_{[a,b]} g\,\mathrm{d}m = \int_a^b f(x)\,\mathrm{d}x.$$

■

R However, an improper Riemann integral does not necessarily has the corresponding Lebesgue integral:

$$f(x) = \sum_{n=1}^{\infty} (-1)^n n \cdot \mathcal{X}_{(1/(n+1),1/n]},\ x \in [0,1]$$

In this case, f is Riemann integrable but not Lebesgue integrable.

12.5. Wednesday for MAT3006

12.5.1. Riemann Integration & Lebesgue Integration

■ **Example 12.6** Compute the integral

$$L = \lim_{n\to\infty}\int_0^1 \frac{nx\log(x)}{1+n^2x^2}\,\mathrm{d}x.$$

Let $f_n(x) = \frac{nx\log(x)}{1+n^2x^2}\mathcal{X}_{(0,1]}$, which is continuous on $[0,1]$, i.e., integrable on $[0,1]$. The goal is to show $L = 0$.

- Note that $f_n(x) \to 0, \forall x \in [0,1]$ pointwisely, as $n \to \infty$.
- Note that $t/(1+t^2) \le \frac{1}{2}, \forall t \ge 0$. Take $t = nx$, we imply

$$|f_n(x)| \le \frac{1}{2}|\log(x)|\mathcal{X}_{(0,1]}$$

We claim that $\frac{1}{2}|\log(x)|\mathcal{X}_{(0,1]} := -\frac{1}{2}\log(x)\mathcal{X}_{(0,1]}$ is integrable: by MCT I,

$$\int -\frac{1}{2}\log(x)\mathcal{X}_{(0,1]}\,\mathrm{d}m = \lim_{n\to\infty}\int_{1/n}^1 -\frac{1}{2}\log(x)\,\mathrm{d}x = \frac{1}{2} < \infty.$$

Therefore, the DCT applies, and

$$\lim_{n\to\infty}\int_{(0,1]}\frac{nx\log(x)}{1+n^2x^2}\,\mathrm{d}x = \int_{(0,1]}\lim_{n\to\infty}\frac{nx\log(x)}{1+n^2x^2}\,\mathrm{d}x = \int_{(0,1]} 0\,\mathrm{d}x = 0$$

However, $f_n(x)$ does not converge to $f(x) \equiv 0$ uniformly on $[0,1]$:

$$\sup_{0\le x\le 1}|f_n(x) - 0| \ge |f_n(1/n) - 0| = \frac{1}{2}\log(n) \to \infty, \text{as } n \to \infty$$

Therefore, we cannot switch integral symbol and limit by using the tools in MAT2006. ■

Proposition 12.6 Suppose that $f(x)$ is a proper Riemann integrable function on $[a,b]$.

Then $f(x)$ is Lebesgue integrable on $[a,b]$ with

$$\int_{[a,b]} f \, \mathrm{d}m = \int_a^b f(x) \, \mathrm{d}x.$$

Proof. Since f is properly Riemann inregrable, we imply $f(x)$ is bounded on $[a,b]$, i.e., $|f(x)| \le K, \forall x \in [a,b]$. Construct the Riemann lower and upper functions with 2^n equal subintervals, denoted as ϕ_n, ψ_n, which follows that

- $\phi_n(x) \le f(x) \le \psi_n(x), \forall n$
- $\phi_n(x)$ is monotone increasing
- $\psi_n(x)$ is monotone decreasing

Now apply bounded convergence theorem on $\psi_n - \phi_n$:

- $|\psi_n(x) - \phi_n(x)| \le 2K$ on $[a,b]$
- $\psi_n - \phi_n \to \psi - \phi$

which implies

$$\begin{aligned}
\int |\psi - \phi| \, \mathrm{d}m &= \int \psi - \phi \, \mathrm{d}m \\
&= \lim_{n\to\infty} \int \psi_n - \phi_n \, \mathrm{d}m = \lim_{n\to\infty} \int \psi_n \, \mathrm{d}m - \lim_{n\to\infty} \int \phi_n \, \mathrm{d}m \\
&= \text{Riemann Upper Sum} - \text{Riemann Lower Sum} \\
&= 0
\end{aligned}$$

Therefore, $\int |\psi - \phi| \, \mathrm{d}m = 0$ implies $\psi(x) = \phi(x)$ a.e. By sandwich theorem,

$$\psi(x) = f(x) = \phi(x) \text{ a.e.}$$

Therefore,

$$\int f \, \mathrm{d}m = \int \phi \, \mathrm{d}m = \lim_{n\to\infty} \int \phi_n \, \mathrm{d}m = \int_a^b f(x) \, \mathrm{d}x$$

where the second equality is by MCT II.

∎

R The improper Riemann integrable functions $f(x)$ is not necessarily Lebesgue integrable. However, if we assume $f(x) \geq 0$, then $f(x)$ is improper Riemann integrable implies $f(x)$ is Lebesgue integrable, with the same integral value.

Proof Outline. Suppose $f(x)$ is improper Riemann integrable on $[a,b]$, where $a,b \in \mathbb{R} \cup \{\pm\infty\}$.

- Construct $f_n = f\mathcal{X}_{[a_n,b_n]}$, with $[a_n,b_n] \subseteq [a_{n+1},b_{n+1}] \subseteq \cdots \subseteq [a,b]$.
- By previous proposition, f_n is proper Riemann integrable implies f_n is Lebesgue integrable.
- Then we apply the MCT I to $\{f_n\}$.

■

12.5.2. Continuous Parameter DCT

Theorem 12.4 — Continous parameter DCT. Let $I, J \subseteq \mathbb{R}$ be intervals, and $f : I \times J \to \mathbb{R}$ be such that

1. for fixed $y \in J$, the function $f(x) := f(x,y)$ is an integrable function over I.
2. for fixed $y \in J$,
$$\lim_{y' \to y} f(x,y') = f(x,y)$$
for almost all $x \in I$
3. There exists integrable $g(x)$ (do not depend on y) such that for all $y \in J$,
$$|f(x,y)| \leq g(x)$$
for almost all $x \in I$.

As a result,
$$F(y) = \int_I f(x,y)\,\mathrm{d}x$$
is a continuous function on J.

R Note that the integrability of $f(x)$ in hypothesis (1) can be weaken into the measurability of $f(x)$: The measurability of $f(x)$ together with hypothesis (3), and DCT implies the integrability of $f(x)$.

Proof. Let $\{y_n\}$ be a sequence on J such that $y_n \to y$. It suffices to show $F(y_n) \to F(y)$. Construct $f_n(x) = f(x, y_n)$, which follows that

- $f_n(x)$ is integrable for all n (by hypothesis (1)) (why check integrable)
- $|f_n(x)| \le g(x)$ a.e. for all n, and $g(x)$ is integrable (by hypothesis (3))
- By hypothesis (2),

$$\lim_{n\to\infty} f_n(x) = f(x,y)$$

Therefore, the DCT applies, and

$$\lim_{n\to\infty}\int_I f_n(x,y_n)\,\mathrm{d}m = \int \lim_{n\to\infty} f_n(x,y_n)\,\mathrm{d}m = \int_I f(x,y)\,\mathrm{d}m$$

Or equivalently,

$$\lim_{n\to\infty} F(y_n) = F(y)$$

■

■ **Example 12.7** Consider $f(x,y) = e^{-x}x^{y-1}$ with $I \times J = (0,\infty)\times[m,M]$, where $0 < m < M < \infty$. We will study the integral

$$\Gamma(y) = \int_0^\infty e^{-x}x^{y-1}\,\mathrm{d}x$$

We check the hypothesis in the Theorem (12.4):

1. For fixed $y \in [m,M]$, $f(x) := f(x,y)$ is indeed measurable on $(0,\infty)$, since $f(x)$ is continous on $(0,\infty)$.
2. The hypothesis (2) follows directly from the contiuity of $f(x,y)$

3.

$$\begin{aligned}|f(x,y)| &\le e^{-x}x^{m-1}\mathcal{X}_{[0,1]} + e^{-x}x^{M-1}\mathcal{X}_{(1,\infty)} \\ &\le x^{m-1}\mathcal{X}_{[0,1]} + e^{-x}x^{M-1}\mathcal{X}_{(1,\infty)}\end{aligned}$$

Here $x^{m-1}\mathcal{X}_{[0,1]}$ is integrable. Following the similar argument in (1), we imply $e^{-x}x^{M-1}\mathcal{X}_{(1,\infty)}$ is integrable as well.

Therefore, $\Gamma(y)$ is continuous for any $m \le y \le M$. Since the choice of $0 < m < M < \infty$ is arbitrary, we imply $T(y)$ is continous on $(0,\infty)$.

In the next lecture we wish to show that

$$F'(y) = \int_I \frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x$$

13.2. Monday for MAT3006

Notations. In this lecture, we let $\int_I f(x,y)\,\mathrm{d}x$ denote the Lebesgue integral.

Theorem 13.3 Let I, J be intervals in $\mathbb{R}$, and $f : I \times J \to \mathbb{R}$ be a function such that

1. For fixed $y \in J$, the function $f(x) := f(x,y)$ is integrable on I
2. $\frac{\partial f}{\partial y}$ exists for any $(x,y) \in I \times J$
3. $\left|\frac{\partial f}{\partial y}(x,y)\right| \le g(x)$ for some integrable function $g(x)$ on I.

Then $F(y) := \int_I f(x,y)\,\mathrm{d}x$ is differentiable on J, with

$$F'(y) = \int_I \frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x$$

Proof. Fix $y \in J$, and consider any sequence $\{y_n\}$ (with $y_n \ne y$) in J converging to y.

Construct the function

$$g_n(x) := \frac{f(x,y_n) - f(x,y)}{y_n - y}$$

which follows that

1. The function g_n is integrable by hypothesis (1)
2. The function $g_n(x)$ converges to $\frac{\partial f}{\partial y}(x,y)$ as $n \to \infty$
3. By MVT, $|g_n(x)| = |\frac{\partial f}{\partial y}(x,\xi)|$, which is bounded by $g(x)$ by hypothesis (3).

Therefore, the DCT applies, and

$$\int_I g_n(x)\,\mathrm{d}x = \frac{1}{y_n - y}\left[\int f(x,y_n)\,\mathrm{d}x - \int f(x,y)\,\mathrm{d}x\right] \to \int_I \frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x$$

In other words, for all sequences $\{y_n\} \to y$ with $y_n \ne y$,

$$\lim_{n\to\infty} \frac{F(y_n) - F(y)}{y_n - y} = \int_I \frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x$$

From the elementary analysis knowledge, in particular, $\lim_{y'\to y} H(y')$ exists (equal to L) if and only if $\lim_{n\to\infty} H(y_n) = L$ for all sequences $\{y_n\} \to y$ with $y_n \ne y$. Therefore,

$$F'(y) := \lim_{y'\to y} \frac{F(y') - F(y)}{y' - y} = \int_I \frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x.$$

■

13.2.1. Double Integral

Definition 13.1 [Measure in $\mathbb{R}^2$] In $\mathbb{R}^2$, we can define the **measure** of the rectangle $A \times B \subseteq \mathbb{R}^2$ with $A, B \in \mathcal{M}$ by

$$m^*(A \times B) = m(A)m(B)$$

In particular, we define

$$x \cdot \infty = \infty \cdot x = (-x) \cdot (-\infty) = \begin{cases} \infty, & \text{if } x > 0 \\ -\infty, & \text{if } x < 0 \\ 0, & \text{if } x = 0 \end{cases}$$

Definition 13.2 [Outer Measure in $\mathbb{R}^2$] Then the outer measure of any $E \subseteq \mathbb{R}^2$ is defined as

$$m^*(E) := \inf \left\{ \sum_{i=1}^{\infty} m(R_i) \middle| E \subseteq \bigcup_{i=1}^{\infty} R_i, R_i = A_i \times B_i, A_i, B_i \in \mathcal{M} \right\}$$

Definition 13.3 [Lebesgue Measurable in $\mathbb{R}^2$] A subset $E \subseteq \mathbb{R}^2$ is Lebesgue measurable if E satisfies the Carathedory Property:

$$m^*(A) = m^*(A \cap E) + m^*(A \setminus E),$$

for any subset $A \subseteq \mathbb{R}^2$.

Product Space of $\mathbb{R}^2$. Given two measurable spaces $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$, in particular, we are interested in

$$(X, \mathcal{A}, \mu) = (Y, \mathcal{B}, \lambda) = (\mathbb{R}, \mathcal{M}, m).$$

Now we want to construct another measurable space in $X \times Y := \mathbb{R}^2$.

1. Start from the "measurable rectangles"

$$\mathcal{A} \times \mathcal{B} = \{A \times B \mid A \in \mathcal{A}, B \in \mathcal{B}\}$$

2. Define the function $\pi : \mathcal{A} \times \mathcal{B} \to [0, \infty]$ by

$$\pi(A \times B) = \mu(A)\lambda(B).$$

3. Let $\mathcal{A} \otimes \mathcal{B}$ be the smallest σ-algebra containing $\mathcal{A} \times \mathcal{B}$. Then by Caratheodory extension theorem, we can extend $\pi : \mathcal{A} \times \mathcal{B} \to [0, \infty]$ to $\tilde{\pi} : \mathcal{A} \otimes \mathcal{B} \to [0, \infty]$ such that

   (a) $(X \times Y, \mathcal{A} \otimes \mathcal{B}, \tilde{\pi})$ is a measurable space

   (b) $\tilde{\pi} \mid_{\mathcal{A} \times \mathcal{B}} = \pi$.

R

- If further we have $\mathcal{A}$ and $\mathcal{B}$ are σ-finite, i.e., there exists $E_i \in \mathcal{A}$ such that $X = \cup_{i=1}^{\infty} E_i$, $\mu(E_i) < \infty, \forall i$, then we can imply the extension $\tilde{\pi}$ is unique.

  (For instance, $\mathbb{R} = \cup_{n \in \mathbb{Z}} [n, n+1]$ and $m([n, n+1]) = 1 < \infty$, i.e., $(\mathbb{R}, \mu, m)$ is σ-finite.)

- Question: we have constructed two measurable space $(\mathbb{R} \times \mathbb{R}, \mathcal{M} \otimes \mathcal{M}, \tilde{\pi})$ and $(\mathbb{R}^2, \mathcal{M}_{\mathbb{R}^2}, m)$. Are they the same?

  Answer : no, but the latter can be obtained from the former by completion process. In particular,

$$m \mid_{\mathcal{M} \otimes \mathcal{M}} = \tilde{\pi}.$$

Let's study the measurable space $(\mathbb{R}\times\mathbb{R}, \mathcal{M}\otimes\mathcal{M}, \pi)$ first, where $f:\mathbb{R}^2\to[-\infty,\infty]$ is a measurable function, i.e., $f^{-1}((a,\infty])\in\mathcal{A}\otimes\mathcal{B}$. In particular, we say $E\subseteq\mathbb{R}\times\mathbb{R}$ is measurable if $E\in\mathcal{M}\otimes\mathcal{M}$ for the moment being (but we will generalize the notion of measurable into $\mathcal{M}_{\mathbb{R}^2}$ in the future).

Definition 13.4 [x-section and y-section] Let $E\subseteq X\times Y$, with $(x,y)\in E$. Define

- the x-section $E_x=\{y\in Y\mid(x,y)\in E\}$, for fixed $x\in X$
- the y-section $E_y=\{x\in X\mid(x,y)\in E\}$, for fixed $y\in Y$.

Proposition 13.2 Suppose that $E\subseteq X\times Y$ is measurable (i.e., $E\in\mathcal{A}\otimes\mathcal{B}$), then $E_x\in\mathcal{B}$ and $E_y\in\mathcal{A}$.

Proof. Construct the set $\mathfrak{A}=\{E\in\mathcal{A}\otimes\mathcal{B}\mid E_x\in\mathcal{B}\}$. It suffices to show $\mathfrak{A}=\mathcal{A}\otimes\mathcal{B}$. We claim that

1. $\mathfrak{A}$ is a σ-algebra
2. $\mathfrak{A}$ contains all $A\times B\in\mathcal{A}\times\mathcal{B}$

If the claim (1) and (2) hold, and since $\mathcal{A}\otimes\mathcal{B}$ is the smallest-σ-algebra containing $\mathcal{A}\times\mathcal{B}$, we imply $\mathcal{A}\otimes\mathcal{B}\subseteq\mathfrak{A}\subseteq\mathcal{A}\otimes\mathcal{B}$, i.e., the proof is complete.

1. (a) Note that $\emptyset\in\mathfrak{A}$, and $X\times Y\in\mathfrak{A}$ since $(X\times Y)_x=Y\in\mathcal{B}$.
   (b) Suppose that $E_i\in\mathfrak{A}$, $i\geq 1$, i.e., $(E_i)_x\in\mathcal{B}$. Observe that
   $$\left(\bigcup_{i=1}^{\infty}E_i\right)_x=\bigcup_{i=1}^{\infty}(E_i)_x\in\mathcal{B},$$
   since $\mathcal{B}$ is a σ-algebra. Therefore, $\cup_{i=1}^{\infty}E_i\in\mathfrak{A}$.
   (c) Suppose that $E\in\mathfrak{A}$, i.e., $(E)_x\in\mathcal{B}$, then
   $$\begin{aligned}(E^c)_x&=\{y\mid(x,y)\in E^c\}\\&=\{y\mid(x,y)\notin E\}\\&=(E_x)^c\in\mathcal{B}\end{aligned}$$

which implies $E^c \in \mathfrak{A}$.

2. For any $A \times B \in \mathcal{A} \times \mathcal{B}$, since $(A \times B)_x = B \in \mathcal{B}$, we imply $(A \times B) \in \mathfrak{A}$.

In conclusion, $\mathfrak{A} = \mathcal{A} \otimes \mathcal{B}$. For all $E \in \mathcal{A} \otimes \mathcal{B}$, we imply $E \in \mathfrak{A}$, i.e., $E_x \in \mathcal{B}$. ■

Proposition 13.3 Sippose that $f : X \times Y \to [-\infty, \infty]$ is measurable. (i.e., $f^{-1}((a, \infty]) \in \mathcal{A} \otimes \mathcal{B}$), then the maps

$$\begin{cases} f_x : & Y \to [-\infty, \infty] \\ \text{with} & f_x(y) := f(x, y) \end{cases}, \qquad \begin{cases} f_y : & X \to [-\infty, \infty] \\ \text{with} & f_y(x) := f(x, y) \end{cases}$$

are measurable. More precisely, $f_x^{-1}((a, \infty]) \in \mathcal{B}$ and $f_y^{-1}((a, \infty]) \in \mathcal{A}$.

Proof.

$$\begin{aligned} f_x^{-1}((a, \infty]) &= \{y \in Y \mid f_x(y) \in (a, \infty]\} \\ &= \{y \in Y \mid f(x, y) > a\} \\ &= \{(u, y) \in X \times Y \mid f(u, y) > a\}_x \\ &= (f^{-1}((a, \infty]))_x \in \mathcal{B} \end{aligned}$$

■

13.5. Wednesday for MAT3006

13.5.1. Fubini's and Tonell's Theorem

Motivation. Given two measurable space $(\mathbb{R}, \mathcal{M}, \mathrm{d}x)$ and $(\mathbb{R}, \mathcal{M}, \mathrm{d}y)$, we have constructed the product measurable space $(\mathbb{R}^2, \mathcal{M} \otimes \mathcal{M}, \mathrm{d}\pi)$. Suppose $f : \mathbb{R}^2 \to [-\infty, \infty]$ is measurable on this space, now we want to show that

$$\int f(x,y)\,\mathrm{d}\pi = \int \left(\int f_y(x)\,\mathrm{d}x \right) \mathrm{d}y = \int \left(\int f_x(y)\,\mathrm{d}y \right) \mathrm{d}x$$

Easier Goal. The proof for the statement above is hard. Consider the easier case where f is a simple function first, i.e., $f(x,y) = \mathcal{X}_E(x,y), E \in \mathcal{M} \otimes \mathcal{M}$, which follows that

$$\begin{aligned}
&\int \mathcal{X}_E(x,y)\,\mathrm{d}\pi = \pi(E) \\
&\int (\mathcal{X}_E)_y(x)\,\mathrm{d}x = \int \mathcal{X}_{E_y}(x)\,\mathrm{d}x = m_X(E_y) \\
&\int (\mathcal{X}_E)_x(y)\,\mathrm{d}y = \int \mathcal{X}_{E_x}(y)\,\mathrm{d}x = m_Y(E_x)
\end{aligned}$$

Therefore, our easier goal is to show that

$$\pi(E) = \int m_X(E_y)\,\mathrm{d}y = \int m_Y(E_x)\,\mathrm{d}x, \quad \forall E \in \mathcal{M} \otimes \mathcal{M}. \tag{13.4}$$

Easiest Goal. Consider the simplest case where $E = A \times B \in \mathcal{M} \otimes \mathcal{M}$, where $A \in \mathcal{M}_X, B \in \mathcal{M}_Y$, which implies

- $\pi(A \times B) = m_X(A) m_Y(B)$

- As shown in the figure (13.8), for fixed $y \in Y$,

$$m_X((A \times B)_y) = \begin{cases} m_X(A), & \text{if } y \in B \\ m_X(\emptyset) = 0, & \text{if } y \notin B \end{cases} = m_X(A)\mathcal{X}_B(y)$$

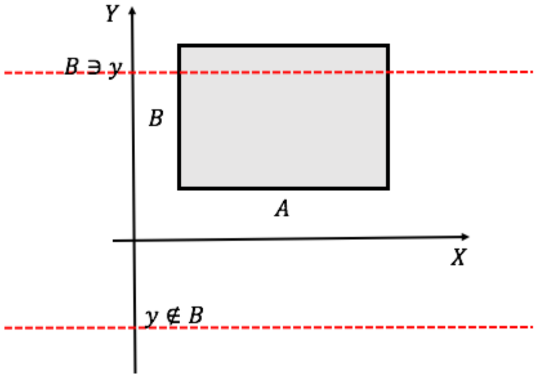


Figure 13.8: Illustration for $m_X((A \times B)_y)$

Therefore, we imply

$$\begin{aligned}\int m_X((A\times B)_y)\,\mathrm{d}y &= \int m_X(A)\mathcal{X}_B(y)\,\mathrm{d}y\\ &= m_X(A)\int \mathcal{X}_B(y)\,\mathrm{d}y\\ &= m_X(A)m_Y(B)\end{aligned}$$

Similarly,

$$\int m_Y((A\times B)_x)\,\mathrm{d}x = m_X(A)m_Y(B).$$

Therefore, the easiest goal (Eq. (13.4)) holds for $E = A \times B \in \mathcal{M} \times \mathcal{M}$.

R Generalization from the easier goal to the real goal is trivial, i.e., applying MCT is ok. The difficulty is that how to show the easier goal (Eq. (13.4)) holds for any $E \in \mathcal{M} \otimes \mathcal{M}$, given that the easier goal (Eq. (13.4)) holds for any $E \in \mathcal{M} \times \mathcal{M}$.

Definition 13.5 [Monotone Class] Let X be a non-empty set. A **monotone class** $\mathcal{T}$ is a collection of subsets of X closed under countable increasing unions and countable decreasing intersections, i.e.,

1. If $E_i \in \mathcal{T}(i \in \mathbb{N})$ and $E_i \subseteq E_{i+1}, \forall i$, then

$$\bigcup_{i=1}^{\infty} E_i \in \mathcal{T}$$

2. If $F_i \in \mathcal{T}(i \in \mathbb{N})$ with $F_i \supseteq F_{i+1}, \forall i$, then

$$\bigcap_{i=1}^{\infty} F_i \in \mathcal{T}$$

R Every σ-algebra is a monotone classs. In particular, for $X = \mathbb{R}$, the collection of subsets $\mathcal{M}$ and $\mathcal{B}$ are both monotone classes.

Definition 13.6 [Smallest Monotone Class] For any $S \subseteq \mathcal{P}(X)$, denote

$$\mathcal{M}(S) := \bigcap_{\mathcal{T} \text{ is a monotone class such that } S \subseteq \mathcal{T}} \mathcal{T},$$

which is also the monotone class. We call $\mathcal{M}(S)$ as the smallest monotone class containing S.

R It's clear that $\mathcal{M}(S) \subseteq \sigma(S)$, where $\sigma(S)$ is the smallest σ-algebra containing S. Question: when do we have $\mathcal{M}(S) = \sigma(S)$?

Theorem 13.5 — **Monotone Class Theorem.** Let X be a non-empty set. If $S \subseteq \mathcal{P}(X)$ is an **algebra** (i.e., $E_1, E_2 \in S \implies E_1 \cup E_2 \in S, E_1 \cap E_2 \in S, E_1^c \in S$), then $\mathcal{M}(S) = \sigma(S)$.

We skip the proof for the monotone class theorem, but you may refer to the proof in the blackboard.

Example 13.3 1. Let $X = \mathbb{R}$, and $S^1 = \{\text{all intervals}\}$ is **not** an algebra, e.g.,

$$[1,2] \in S^1 \implies [1,2]^c = (-\infty, 1) \cup (2, \infty) \notin S^1.$$

However, $S = \{\text{finite disjoint union of intervals}\}$ is an algebra. Therefore,

$$\mathcal{M}(S) = \sigma(S) := \mathcal{B}(\text{Borel } \sigma\text{-algebra}).$$

2. Let $X = \mathbb{R}^2$, and define

$$S = \left\{ \text{finite disjoint union of measurable rectangles } \bigcup_{i=1}^{k} (A_i \times B_i) \,\middle|\, A_i, B_i \in \mathcal{M} \right\}$$

Then S is an algebra, for instance, as shown in the Fig. (13.9), $(A \times B)^c = (A^c \times \mathbb{R}) \cup (A \times B^c)$ is a disjoint union of 2 measurable rectangles.

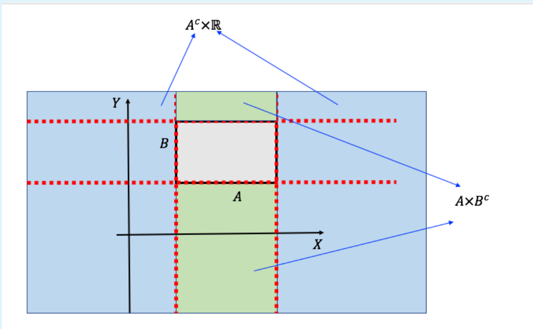


Figure 13.9: Illustration for $(A \times B)^c$

Therefore, $\mathcal{M}(S) = \sigma(S) := \mathcal{M} \otimes \mathcal{M}$

Proposition 13.10 For all $E \in \mathcal{M} \otimes \mathcal{M}$, we have

$$\pi(E) = \int m_Y(E_x)\,\mathrm{d}x = \int m_X(E_y)\,\mathrm{d}y \tag{13.5}$$

Proof. Construct

$$\mathcal{A} = \left\{ E \in \mathcal{M} \otimes \mathcal{M} \,\middle|\, \begin{array}{l} x \mapsto m_Y(E_x) \text{ is a measurable function of } x \\ y \mapsto m_X(E_y) \text{ is a measurable function of } y \\ \text{Eq. (13.5) holds} \end{array} \right\}$$

- Claim 1: $\mathcal{A}$ is a monotone class

- Claim 2: Any finite disjoint union of measurable rectangles is in $\mathcal{A}$:

$$\bigcup_{i=1}^{k}(A_i \times B_i) \in \mathcal{A}, \quad k \in \mathbb{N}$$

If claim (1),(2) holds, then $S \subseteq \mathcal{A}$, where

$$S = \{\text{finite disjoint union of measurable rectangles}\}$$

which follows that

$$\mathcal{M}(S) \subseteq \mathcal{A}.$$

By monotone class theorem, $\sigma(S) = \mathcal{M}(S) \subseteq \mathcal{A}$, i.e.,

$$\mathcal{M} \otimes \mathcal{M} = \sigma(S) = \mathcal{M}(S) \subseteq \mathcal{A} \subseteq \mathcal{M} \otimes \mathcal{M} \implies \mathcal{M} \otimes \mathcal{M} = \mathcal{A}.$$

Therefore, (13.5) holds for all $E \in \mathcal{A} = \mathcal{M} \otimes \mathcal{M}$.

We left the proof for claim (1) in next class. Now we give a proof for claim (2):

- For any $E = \cup_{i=1}^{k}(A_i \times B_i)$,

  $$m_Y(E_x) = \sum_{i=1}^{k} m_Y(B_i)\mathcal{X}_{A_i}(x)$$

  is a simple function on x, and therefore measurable.

- Similarly,

  $$m_X(E_y) = \sum_{i=1}^{k} m_X(A_i)\mathcal{X}_{B_i}(y)$$

  is also measurable.

- By the easiest goal, (13.5) also holds.

Therefore, claim (2) is true. ■

14.2. Monday for MAT3006

14.2.1. Tonelli's and Fubini's Theorem

Proposition 14.1 For all $E \in \mathcal{M} \otimes \mathcal{M}$,we have

$$\int m_Y(E_x)\,\mathrm{d}x = \int m_X(E_y)\,\mathrm{d}y = \pi(E), \tag{14.1}$$

where $\pi(\cdot)$ is a measure on $\mathcal{M} \otimes \mathcal{M}$.

Here note that

$$\begin{aligned} m_X(E_y) &:= \int (\mathcal{X}_E)_y(x)\,\mathrm{d}x \\ m_Y(E_x) &:= \int (\mathcal{X}_E)_x(y)\,\mathrm{d}y \end{aligned}$$

Proof. Construct

$$\mathcal{A} = \left\{ E \in \mathcal{M} \otimes \mathcal{M} \;\middle|\; \begin{array}{l} x \mapsto m_Y(E_x) \text{ measurable} \\ y \mapsto m_X(E_y) \text{ measurable} \\ \text{(14.1) holds for } E \end{array} \right\}$$

Following the proof given in the last lecture, it suffices to show $\mathcal{A}$ is a monotone class:

- Construct

  $$\mathcal{A}_k = \mathcal{A} \cap \{E \in \mathcal{M} \otimes \mathcal{M} \mid E \subseteq [-k,k] \times [-k,k]\}.$$

  We first show that $\mathcal{A}_k$ is a monotone class for all $k \in \mathbb{N}$:

  1. Suppose that $E_n \subseteq E_{n+1}, \forall n$ and $E_n \in \mathcal{A}_k$, and we aim to show $E := \cup_{n=1}^{\infty} E_n \in \mathcal{A}_k$. Consider the function $f_n(x) = m_Y((E_n)_x)$, which is measurable for all n, and $f_n(x) \le f_{n+1}(x)$ for all n, since $E_n \subseteq E_{n+1}$.

     The MCT I implies that $f(x) = m_Y(E_x)$ is measurable with

     $$\int m_Y(E_x)\,\mathrm{d}x = \lim_{n\to\infty} \int m_Y((E_n)_x)\,\mathrm{d}x \overset{(a)}{=} \lim_{n\to\infty} \pi(E_n) \overset{(b)}{=} \pi(E)$$

where (a) is because that $E_n \in \mathcal{A}$; and (b) is due to the exercise in Hw3. Similarly, $y \mapsto m_X(E_y)$ is measurable, with $\int m_X(E_y)\,\mathrm{d}y = \pi(E)$. Therefore, $E \in \mathcal{A}$, i.e., $E \in \mathcal{A}_k$ as well.

2. Suppose that $F_i \in \mathcal{A}_k, F_i \supseteq F_{i+1}$, and we aim to show $F := \cap_{i=1}^{\infty} F_i \in \mathcal{A}_k$. Construct the measurable function $g_n(x) = m_Y((F_n)_x)$, and $g_n(x) \ge g_{n+1}(x)$; $|g_n(x)| \le g_1(x)$, with $g_1(x)$ integrable. (You may see the bounded rectangle in $\mathcal{A}_k$ matters here)

   The DCT implies that $g(x) = m_Y(F_x)$ is measurable, with

$$\int m_Y(F_x)\,\mathrm{d}x = \lim_{n\to\infty} \int g_n\,\mathrm{d}x = \lim_{n\to\infty} \pi(F_n) = \pi(F).$$

   Similarly, $y \mapsto m_X(F_y)$ is measurable, with $\int m_X(F_y)\,\mathrm{d}y = \pi(F)$. Therefore, $F \in \mathcal{A}_k$.

Together with the results from last lecture, we conclude that claim (1) and (2) holds for $\mathcal{A}_k$. Following the similar idea of the results obtained from last lecture, we conclude that $\mathcal{A}_k = \{E \in \mathcal{M} \otimes \mathcal{M} \mid E \subseteq [-k,k] \times [-k,k]\}$.

- Then we show $\mathcal{A}$ is a monotone class, i.e., closed under countable decreasing intersections.

  Suppose that $F_i \in \mathcal{A}, F_i \supseteq F_{i+1}$, we aim to show that $F := \cap F_i \in \mathcal{A}$.

  Construct

$$F_i^{(k)} = F_i \cap ([-k,k] \times [-k,k]),$$

  which follows that $F_i^{(k)} \supseteq F_{i+1}^{(k)}$, and $F_i^{(k)} \in \mathcal{A}_k$ since $F_i^{(k)} \in \mathcal{M} \otimes \mathcal{M}$ and $F_i^{(k)} \subseteq [-k,k] \times [-k,k]$. We denote $F^{(k)} = \cap_{i=1}^{\infty} F_i^{(k)}$. The previous result implies that $F^{(k)} \in \mathcal{A}_k$, i.e.,

$$\int m_Y((F^{(k)})_x)\,\mathrm{d}x = \pi(F^{(k)})$$

  Now note that $F^{(1)} \subseteq F^{(2)} \subseteq \cdots$, and $F = \cup_{k\in\mathbb{N}} F^{(k)}$. Therefore, applying MCT gives

$$\int m_Y(F_x)\,\mathrm{d}x = \lim_{k\to\infty} \int m_Y((F^{(k)})_x)\,\mathrm{d}x = \lim_{k\to\infty} \pi(F^{(k)}) = \pi(F).$$

Therefore, F satisfies (14.1), i.e., $F \in \mathcal{A}$

■

Theorem 14.3 — **Tonelli's Theorem.** Let $F : \mathbb{R}^2 \to [0,\infty]$ be measurable under the space $(\mathbb{R}^2, \mathcal{M} \otimes \mathcal{M}, \pi)$. Then

$$\begin{cases} x \mapsto \int F(x,y)\,\mathrm{d}y \\ y \mapsto \int F(x,y)\,\mathrm{d}x \end{cases} \text{is measurable,}$$

and

$$\int F\,\mathrm{d}\pi = \int \left(\int F(x,y)\,\mathrm{d}x \right) \mathrm{d}y = \int \left(\int F(x,y)\,\mathrm{d}y \right) \mathrm{d}x$$

Proof. Let

$$\phi_n(x,y) = \sum_{k=0}^{4^n} (k \cdot 2^{-n}) \mathcal{X}_{F^{-1}([k\cdot 2^{-n},(k+1)\cdot 2^{-n}])} + 2^n \mathcal{X}_{F^{-1}(2^n,\infty]}$$

We just re-write the terms above as $\sum_k \alpha_k \mathcal{X}_{E_k}$. Our constructed $\phi_n(x,y)$ is a monotone increasing simple function such that $\phi_n \to F$ pointwise. It follows that

$$\int F\,\mathrm{d}\pi = \lim_{n\to\infty} \int \phi_n\,\mathrm{d}\pi \tag{14.2a}$$

$$= \lim_{n\to\infty} \int \left(\sum_k \alpha_k \mathcal{X}_{E_k} \right) \mathrm{d}\pi \tag{14.2b}$$

$$= \lim_{n\to\infty} \sum_k \alpha_k \int \mathcal{X}_{E_k}\,\mathrm{d}\pi = \lim_{n\to\infty} \sum_k \alpha_k \pi(E_k) \tag{14.2c}$$

$$= \lim_{n\to\infty} \sum_k \alpha_k \int \left(\int \mathcal{X}_{E_k}(x,y)\,\mathrm{d}x \right) \mathrm{d}y \tag{14.2d}$$

$$= \lim_{n\to\infty} \int \int \left(\sum_k \alpha_k \mathcal{X}_{E_k}(x,y) \right) \mathrm{d}x\,\mathrm{d}y \tag{14.2e}$$

$$= \lim_{n\to\infty} \int \left(\int \phi_n(x,y)\,\mathrm{d}x \right) \mathrm{d}y \tag{14.2f}$$

$$= \int \lim_{n\to\infty} \left(\int \phi_n(x,y)\,\mathrm{d}x \right) \mathrm{d}y \tag{14.2g}$$

$$= \int \int \lim_{n\to\infty} \phi_n(x,y)\,\mathrm{d}x\,\mathrm{d}y \tag{14.2h}$$

$$= \int \int F(x,y)\,\mathrm{d}x\,\mathrm{d}y \tag{14.2i}$$

where (14.2a) is by the MCT I on ϕ_n; (14.2c) is by the linearity of integral; (14.2d) is by proposition (14.1) (14.2e) is by the linearity of integral; (14.2g) is by the MCT I on $f_n(y) = \int \phi_n(x,y)\,\mathrm{d}x$; (14.2h) is by the MCT I on $g_n(x) = \phi_n(x,y)$; (14.2i) is because that $\phi_n(x,y) \to F(x,y)$. ■

Theorem 14.4 — **Fubini's Theorem.** Suppose that $F : \mathbb{R}^2 \to [-\infty,\infty]$ is integrable, then

$$\int F\,\mathrm{d}\pi = \int \left(\int F(x,y)\,\mathrm{d}x \right) \mathrm{d}y = \int \left(\int F(x,y)\,\mathrm{d}y \right) \mathrm{d}x$$

Proof. Suppose $F = F^+ - F^-$, where $F^\pm$ are both integrable. Applying Tonell's theorem on both F^- and F^+ and the linearity of integrals gives the desired result. ■

15.2. Monday for MAT3006

15.2.1. Applications on the Tonell's and Fubini's Theorem

Theorem 15.2 — Tonell. Let $f: \mathbb{R}^2 \to [0,\infty]$ be a measurable function (i.e., $f^{-1}((a,\infty]) \in \mathcal{M} \otimes \mathcal{M}$), then

$$\int f \, d\pi = \int \left(\int f(x,y) \, dx \right) dy = \int \left(\int f(x,y) \, dy \right) dx$$

Theorem 15.3 — Fubini. Let $f: \mathbb{R}^2 \to [-\infty,\infty]$ be integrable (i.e., $f = f^+ - f^-$ with $f^\pm : \mathbb{R}^2 \to [0,\infty]$ measurable and $\int f^\pm \, dx < \infty$), then

$$\int f \, d\pi = \int \left(\int f(x,y) \, dx \right) dy = \int \left(\int f(x,y) \, dy \right) dx$$

Corollary 15.2 Suppose that $f: \mathbb{R}^2 \to [-\infty,\infty]$ is measurable, and either

$$\int \left(\int |f(x,y)| \, dx \right) dy \tag{15.2a}$$

or

$$\int \left(\int |f(x,y)| \, dy \right) dx \tag{15.2b}$$

exists, then f is integrable, and the result of Fubini follows. (i.e., one can switch the order of integration as long as the integral of $|f|$ exists)

Proof. If (15.2a) or (15.2b) exists (is finite), then Tonell's Theorem implies that $|f|$ is integrable, which implies f is integrable.

Therefore, the assumption of Fubini's theorem holds, and the proof is comptete. ■

R The advantage for corollary (15.2) is that computing (15.2a) or (15.2b) is easier than showing the integrability of f in general.

■ **Example 15.2** Compute the double integral

$$\int_0^1 \int_0^x \sqrt{\frac{1-y}{x-y}}\,\mathrm{d}y\,\mathrm{d}x.$$

Construct the function $f(x,y) := \sqrt{\frac{1-y}{x-y}}\mathcal{X}_E(x,y)$, with E shown in Fig. (15.1).

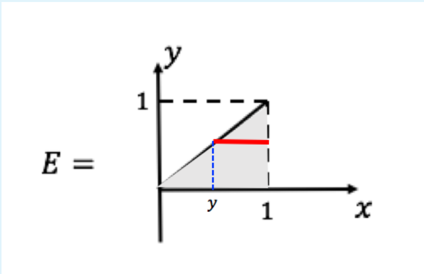


Figure 15.1: Illustration for integral domain E

We want to compute $\int f(x,y)\,\mathrm{d}\pi$ and show that

$$\int_0^1 \int_0^x \sqrt{\frac{1-y}{x-y}}\,\mathrm{d}y\,\mathrm{d}x = \int f(x,y)\,\mathrm{d}\pi.$$

- Consider the integral

$$\begin{aligned}
\int \left(\int f(x,y)\,\mathrm{d}x\right)\mathrm{d}y &= \int_0^1 \left(\int_y^1 \sqrt{\frac{1-y}{x-y}}\,\mathrm{d}x\right)\mathrm{d}y && (15.3a)\\
&= \int_0^1 \sqrt{1-y}\left(\int_y^1 \frac{1}{\sqrt{x-y}}\,\mathrm{d}x\right)\mathrm{d}y && (15.3b)\\
&= \int_0^1 \sqrt{1-y}\left(\int_0^{1-y} \frac{1}{\sqrt{t}}\,\mathrm{d}t\right)\mathrm{d}y && (15.3c)\\
&= \int_0^1 \sqrt{1-y}\cdot(2\sqrt{1-y})\,\mathrm{d}y && (15.3d)\\
&= 2\int_0^1 (1-y)\,\mathrm{d}y && (15.3e)\\
&= 1 && (15.3f)
\end{aligned}$$

where the justification of (15.4a) is from Fig. (15.1).

- Therefore, $\int(\int|f(x,y)|\,\mathrm{d}x)\,\mathrm{d}y<\infty$. Mreover, f is continous on E°, i.e., measurable on E° (it's clear that a continous function is measurable). Since ∂E is null, we imply f is measurable on $E:=E^\circ\cup\partial E$.

- Therefore, the assumption of Corollary (15.2) holds, and we imply that

$$\int\left(\int f(x,y)\,\mathrm{d}y\right)\mathrm{d}x=\int\left(\int f(x,y)\,\mathrm{d}x\right)\mathrm{d}y$$

It's clear that

$$\int_0^1\int_0^x\sqrt{\frac{1-y}{x-y}}\,\mathrm{d}y\,\mathrm{d}x=\int\left(\int f(x,y)\,\mathrm{d}y\right)\mathrm{d}x,$$

and therefore

$$\int_0^1\int_0^x\sqrt{\frac{1-y}{x-y}}\,\mathrm{d}y\,\mathrm{d}x=1.$$

Process of Completion. We have two measures on $\mathbb{R}^2$:

- $\mathcal{M}\otimes\mathcal{M}$, and

- $\mathcal{M}_{\mathbb{R}^2}$, given by

$$\mathcal{M}_{\mathbb{R}^2}=\{E\subseteq\mathbb{R}^2\mid m^*(A)=m^*(A\cap E)+m^*(A\cap E^c)\text{ for all subsets }A\subseteq\mathbb{R}^2\}$$

Here $\mathcal{M}_{\mathbb{R}^2}$ equals the completion of $\mathcal{M}\otimes\mathcal{M}$, i.e., all $E\subseteq\mathcal{M}_{\mathbb{R}^2}$ can be decomposed as

$$E=B\cup(E\setminus B),$$

where $B\in\mathcal{M}\otimes\mathcal{M}$ and $E\setminus B\in\mathcal{M}_{\mathbb{R}^2}$ with $\pi(E\setminus B)=0$.

Question: does Tonell's theorem holds for (Lebesgue) measurable functions $f:\mathbb{R}^2\to[0,\infty]$ (i.e., $f^{-1}((a,\infty])\in\mathcal{M}_{\mathbb{R}^2}$ for any $a\in[0,\infty)$?)

Answer: Yes. To see so, we just need the following proposition

Proposition 15.4 Let $(\mathbb{R}^2,\mathcal{M}_{\mathbb{R}^2},\pi)$ be the Lebesgue measure on $\mathbb{R}^2$, and $N\in\mathcal{M}_{\mathbb{R}^2}$ be such that $\pi(N) = 0$. Then for almost all values of $x \in \mathbb{R}$, $N_x \in \mathcal{M}$ and $m_Y(N_x) = 0$.

Proof. For $N \in \mathcal{M}_{\mathbb{R}^2}$. By hw3, there exists $B' \in \mathcal{M} \otimes \mathcal{M}$ such that $N \subseteq B'$, with

$$\pi(B') = \pi(N).$$

If N is null, then $\pi(B') = 0$. By Tonell's theorem on $M \otimes \mathcal{M}$, we imply

$$\pi(B') = \int m_Y(B'_x)\,\mathrm{d}x = \int m_X(B'_y)\,\mathrm{d}y = 0$$

Therefore, $m_Y(B'_x) = 0$ for almost all $x \in \mathbb{R}$. Since $N \subseteq B'$, we imply $N_x \subseteq B'_x$, i.e., N_x is also a null set. Therefore, $N_x \in \mathcal{M}$ and $m_Y(N_x) = 0$. ■

■ **Example 15.3** Consider the integral

$$\int_0^\infty \int_0^\infty y e^{-y^2(1+x^2)}\,\mathrm{d}y\,\mathrm{d}x$$

Define $f(x,y) = ye^{-y^2(1+x^2)}$, which is continous on $(0,\infty)\times(0,\infty)$, and therefore measurable. It follows that

$$\begin{aligned}
\int_0^\infty \int_0^\infty f(x,y)\,\mathrm{d}y\,\mathrm{d}x &= \int_0^\infty \left(\lim_{n\to\infty}\int_0^n f(x,y)\,\mathrm{d}y\right)\mathrm{d}x && (15.4a)\\
&= \int_0^\infty \left(\frac{1}{1+x^2}\frac{1}{2}\right)\mathrm{d}x && (15.4b)\\
&= \lim_{n\to\infty}\int_0^n \frac{1}{2}\frac{1}{1+x^2}\,\mathrm{d}x && (15.4c)\\
&= \frac{\pi}{4} && (15.4d)
\end{aligned}$$

where (15.4a) is by applying MCT I on the function $f(x,y)\mathcal{X}_{[0,n]}$; (15.4b) and (15.4d) is by computation; (15.4c) is by applying MCT I on the function $\frac{1}{1+x^2}\frac{1}{2}\mathcal{X}_{[0,n]}$.

By corollary (15.2),

$$\int_0^\infty \int_0^\infty y e^{-y^2(1+x^2)}\,\mathrm{d}x\,\mathrm{d}y = \frac{\pi}{4}$$

Or equivalently,

$$\int_0^\infty ye^{-y^2}\int_0^\infty e^{-x^2y^2}\,\mathrm{d}x\,\mathrm{d}y = \frac{\pi}{4}$$

By applying MCT I on $e^{-x^2y^2}\mathcal{X}_{[0,n]}$, we have

$$\int_0^\infty ye^{-y^2}\lim_{n\to\infty}\int_0^n e^{-x^2y^2}\,\mathrm{d}x\,\mathrm{d}y = \frac{\pi}{4}$$

By change of variable with $t = xy$, we imply

$$\int_0^\infty ye^{-y^2}\lim_{n\to\infty}\frac{1}{y}\int_0^{ny} e^{-t^2}\,\mathrm{d}t\,\mathrm{d}y = \frac{\pi}{4}$$

Or equivalently,

$$\int_0^\infty e^{-y^2}\int_0^\infty e^{-t^2}\,\mathrm{d}t\,\mathrm{d}y = \frac{\pi}{4}$$

Therefore, we conclude that

$$\left(\int_0^\infty e^{-y^2}\,\mathrm{d}y\right)^2 = \frac{\pi}{4} \implies \int_0^\infty e^{-x^2}\,\mathrm{d}x = \frac{\sqrt{\pi}}{2}$$

■