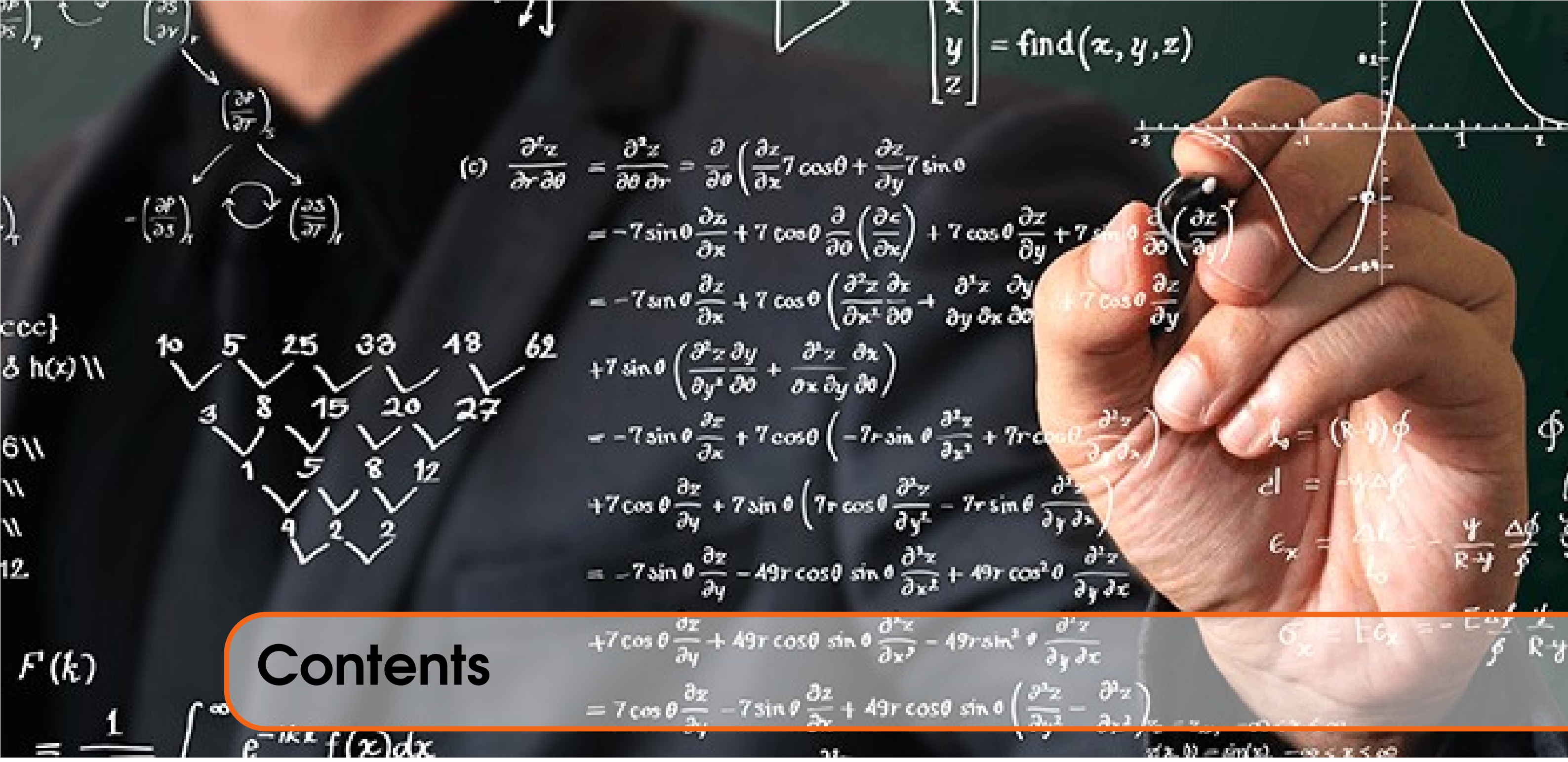
PMATH 450 Course Notes

University of Waterloo

The One And Only
Waterloo 76er
Bill Zhuo



Contents

I Lebesgue Integration

1 Riemann Integration 7

1.1 Axiom of Choice 7

1.2 Review of the Riemann Integral 9

1.2.1 Intuition for Lebesgue Integral 11

1.3 Ideal Notion of a Measure Function 12

1.4 Construction of Lebesgue Outer Measure 14

1.5 Construction of Lebesgue Measure 16

1.5.1 Lebesgue Measurable Sets and Lebesgue Measure 16

1.5.2 Lebesgue's Characterization of Riemann Integrability 21

2 Measurable Functions 27

2.1 Definitions 27

2.2 Properties of Measurable Functions 29

2.2.1 Almost Everywhere 31

3 Lebesgue Integral 33

3.1 Definition of the Lebesgue Integral 33

3.2 Properties of Lebesgue Integral 35

4 MCT and DCT 41

4.1 Monotone Convergence Theorem and Fatou's Lemma 41

4.5 Dominated Convergence Theorem 46

4.6 Relationship Between Riemann and Lebesgue Integrals 47

5 L^p Spaces . . . 49

6 Hölder's Inequality . . . 53
6.1 Hölder's Inequality 53
6.2 Minkowski's Inequality 55

7 Riesz-Fischer Theorem . . . 57

8 Lusin's Theorem and Fubini's Theorem . . . 59
8.1 Statement of Lusin's Theorem 59
8.4 Fubini's Theorem 64

9 Hilbert Spaces . . . 65
9.1 Inner Products and Orthogonality 65
9.3 Bessel's Inequality 67
9.4 Bases for Hilbert Spaces 68

II Fourier Analysis

10 Fourier Analysis on the Circle . . . 75
10.1 History 75
10.2 Basic Notation 75
10.3 Fourier Series and Fourier Coefficients 76
10.4 Parseval's Theorem 77
10.5 More About the Fourier Coefficients of L^1 Functions 79
10.6 Basic Properties of the Fourier Coefficients 79

11 The Dirichlet Kernel . . . 83
11.1 Dirichlet Kernel 83
11.3 Intuitive Idea Why We Might Hope For $S_N f(x) \to f(x)$ 84

12 Failure of Convergence of Fourier Series . . . 87
12.1 Functional Analysis Ideas 87
12.2 Application - An L^1 Function with a Divergent Fourier Series 89
12.3 Application - A continuous function with a Divergent Fourier Series 90

Lebesgue Integration

1 Riemann Integration 7
1.1 Axiom of Choice
1.2 Review of the Riemann Integral
1.3 Ideal Notion of a Measure Function
1.4 Construction of Lebesgue Outer Measure
1.5 Construction of Lebesgue Measure

2 Measurable Functions 27
2.1 Definitions
2.2 Properties of Measurable Functions

3 Lebesgue Integral 33
3.1 Definition of the Lebesgue Integral
3.2 Properties of Lebesgue Integral

4 MCT and DCT 41
4.1 Monotone Convergence Theorem and Fatou's Lemma
4.5 Dominated Convergence Theorem
4.6 Relationship Between Riemann and Lebesgue Integrals

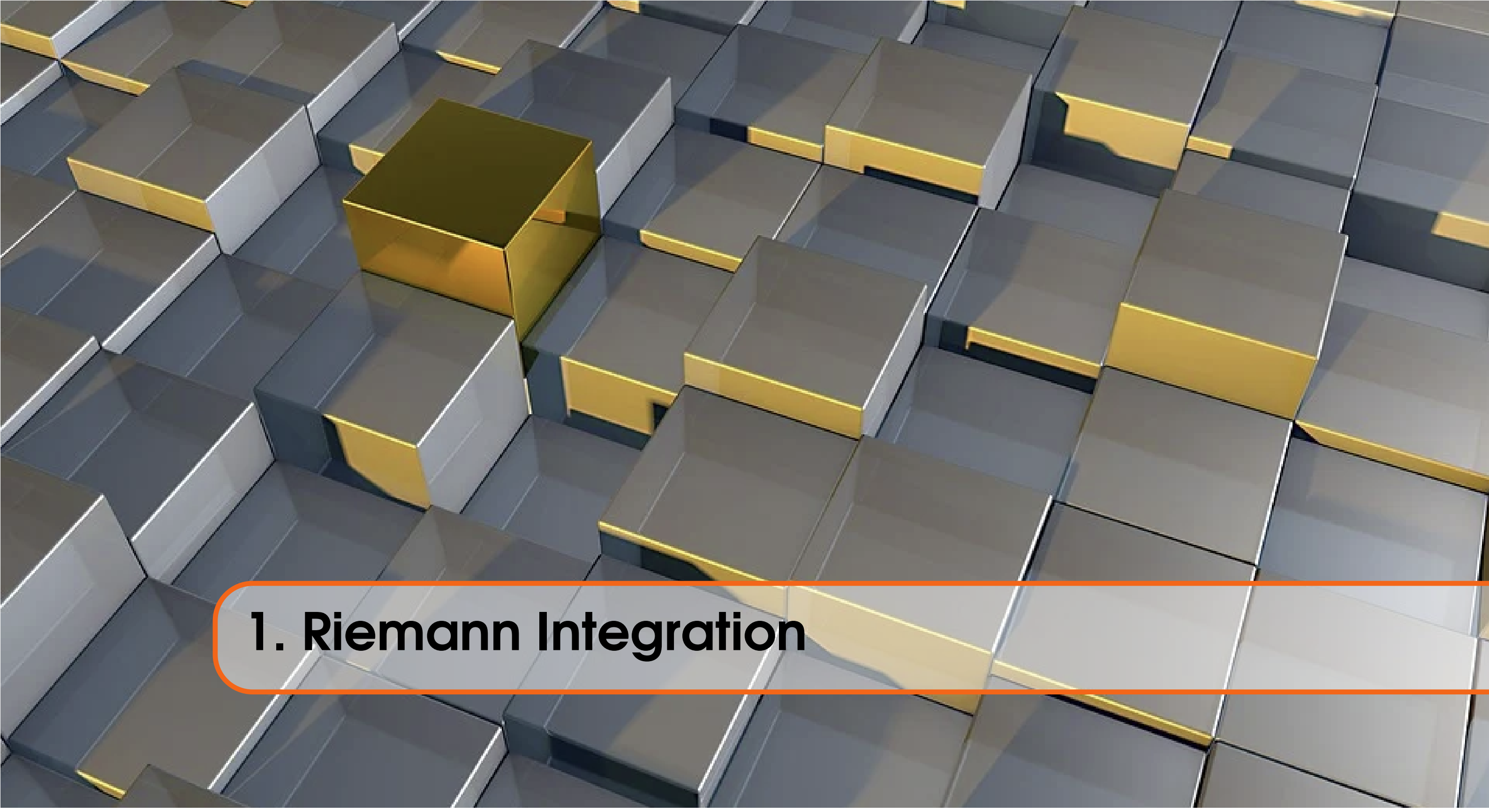
5 L^p Spaces 49

6 Hölder's Inequality 53
6.1 Hölder's Inequality
6.2 Minkowski's Inequality

7 Riesz-Fischer Theorem 57

8 Lusin's Theorem and Fubini's Theorem . 59
8.1 Statement of Lusin's Theorem
8.4 Fubini's Theorem

9 Hilbert Spaces 65
9.1 Inner Products and Orthogonality
9.3 Bessel's Inequality
9.4 Bases for Hilbert Spaces



1. Riemann Integration

1.1 Axiom of Choice

Mathematics is built on a rigorous foundation in which rules (or axioms) are specified. One of these, which might seem quite mundane, has profound consequences - **the Axiom of Choice**.

Axiom 1 — Axiom of Choice. Let $\mathscr{F}$ be a non-empty collection of non-empty sets, say

$$\mathscr{F} = \{A_\lambda : \lambda \in \Lambda\}$$

where the index set Λ and each set A_λ is non-empty. Then, there is a function

$$f : \Lambda \to \bigcup_{\lambda \in \Lambda} A_\lambda$$

such that $f(\lambda) \in A_\lambda$ for each $\lambda \in A$.

R This axiom means that we can pick one element from each set in the collection. Λ can be finite, countable, or even uncountable.

Definition 1.1.1 — Partial Order and Total Order. Let S be a set. A relation $\leq$ on S is called a **partial order** if

1. **Reflexive:** $x \leq x$ for all $x \in S$
2. **Anti-symmetric:** $x \leq y$ and $y \leq x$ implies $x = y$ for all $x, y, z \in S$
3. **Transitive:** $x \leq y$ and $y \leq z$ implies $x \leq z$ for all $x, y, z \in S$

A partial order on S is called a total order if for every $x, y \in S$ either $x \leq y$ or $y \leq x$.

■ **Example 1.1**

1. The usual notion of $\leq$ on $\mathbb{R}$ is a total order
2. **Partial Order by Inclusion:** Given a set S, let $X = \mathscr{P}(S)$ (all the subsets) and define $\leq$ on X by the rule $A \leq B$ if $A \subseteq B$. We will refer to this as the partial order by inclusion.

■

Exercise 1.1 — Partial order by Inclusion is not a total order. This is obvious. Consider the set $\{0,1,2,3\}$ with partial order by inclusion. And clearly we cannot compare $\{0,1\}$ and $\{0,2\}$. There cannot be a total order. ■

Definition 1.1.2 — Well-Ordering. A partial ordered set $(S,\leq)$ is said to be **well-ordered** if every non-empty subset of S has a smallest element. In other words,

If $T \subseteq S$ and T $\emptyset$, then there exists $x \in T$ such that every $y \in T$ satisfies $x \leq y$.

R Any well-ordered set is totally ordered since within every two-element set, we can tell which one is "larger".

■ **Example 1.2**
1. $\mathbb{N}$ is well ordered
2. $\mathbb{Z}$ is totally ordered but not well-ordered since we cannot find the smallest element in the set
3. $\mathbb{Q}^+ = \{r \in \mathbb{Q} : r \geq 0\}$ is not well-ordered since $\mathbb{Q}^+ = \{r \in \mathbb{Q} : r > 0\}$ has no minimal element.

■

Exercise 1.2 Show that every countable set can be well-ordered. Recall: countable means there is a bijection with $\mathbb{N}$.
Solution:
Let A be any countable set. Then, there exists a bijection $f : A \to \mathbb{N}$. For any two $x,y \in A$, there exists $f(x), f(y) \in \mathbb{N}$. We define $\leq$ to be $x \leq y$ if $f(x) \leq f(y)$ in the usual order in $\mathbb{N}$. This is clearly a partial order. Then, for any subset of A, we can use f to find the minimal element in the set since $\mathbb{N}$ is well-ordered. ■

Theorem 1 — Well Ordering Principle-Derived from AOC. Every set can be well ordered. In other words, on every set it is possible to define a partial order that makes the set well ordered.

Definition 1.1.3 — Upper Bound, Maximal, Chain. Let $(S,\leq)$ be poset. An upper bound for $T \subseteq S$ is an element $s \in S$ such that $s \geq t$ for every $t \in T$.
An element $s \in S$ is said to be maximal if whenever $x \in S$ and $x \geq s$, then $x = s$.
A chain is a totally ordered subset of S.

R
1. Upper bound is not necessarily in T
2. A poset does not need to have a maximal element nor unique. For example, $\mathbb{N}$ does not have a maximal element. Or a finite example, $\{a,b,c\}$ with $a \leq b, a \leq c$ and $x \leq x$ for $x = a,b,c$. But b and c are not related. In fact, b and c are maximal.

■ **Example 1.3** Consider $\mathbb{R}^2$ with the order given in example 1.1.2, the set $\{(x,x) : x \in \mathbb{R}\}$ is a chain. (it's like a line with orientation) ■

■ **Example 1.4** Let S be any non-empty set. Let $X = \mathscr{P}(S)$ and give X the partial order by inclusion. Then S is an upper bound for X. Note that $S \in X$ and that if $B \in X$, then clearly $B \subseteq S$. **It is also unique.** ■

Theorem 2 — Zorn's Lemma-Derived from AOC. Suppose S is a non-empty poset in which each chain has an upper bound. Then S has a maximal element.

Theorem 3 — AOC Equivalence. TFAE:

1. AOC
2. Zorn's Lemma
3. Well-Ordering Principle.

Theorem 4 Every vector space has a basis.

Proof. Let V be a vector space and S be the set of any linearly independent subsets of V partially ordered by inclusion. Let C be any chain in S. (The existence is clear.)
Let $Y = \cup_{W \in C} W$. Clearly, $Y \supseteq W$ for all $W \in C$. We claim $Y \in S$.
Let $x_1, x_2, \ldots, x_n \in Y$, then for each $i \in \{1, \ldots, n\}$, there is some $W_i \in C$ such that $x_i \in W_i$. Since there are only finitely many of them, we have one of them being maximal containing $\{x_1, x_2, \ldots, x_n\}$ in $C \subseteq S$, say it's W^*. Note that W^* is linearly independent set since it's in S. Thus,

$$\sum_{i=1}^{n} \alpha_i x_i = 0 \Longrightarrow a_i = 0, \forall i$$

That means Y is a linearly independent subset of V, thus in S. Now, we have all the conditions required to use Zorn's Lemma. This means we can find a maximal in S.
Let M be a maximal element of S, we claim it is a basis for V. (Linearly independence is clear) Suppose it is not a spanning set. Then, there must exist some $x \in V$ such that $x \notin \operatorname{span}(M)$ but $M \cup \{x\}$ is still a linearly independent set, so it is in S, and it properly contains M, thus, $M \cup \{x\} \geq M$ but not equal. But this contradicts the assumption that M was a maximal element of S. Thus, M is a spanning set.
Thus, we are done. ■

1.2 Review of the Riemann Integral

Definition 1.2.1 — Riemann Sums and Refinement. Let $f : [a,b] \to \mathbb{R}$ be bounded. Consider a partition $P : a = x_0 < x_1 < \cdots < x_n = b$ of $[a,b]$. From the upper the lower Riemann sums:

$$U(f,P) = \sum_{i=1}^{n} \sup f\big|_{[x_{i-1},x_i]} (x_i - x_{i-1})$$
$$L(f,P) = \sum_{i=1}^{n} \inf f\big|_{[x_{i-1},x_i]} (x_i - x_{i-1})$$

Note that $U(f,P) \geq L(f,P), \forall f, P$. We say the parition Q is a refinement of P if Q contains P and possibly more points. Then, $U(f,Q) \leq U(f,P)$ and $L(f,Q) \geq L(f,P)$. We can also have common refinement being the union of two partitions P_1, P_2, then

$$U(f,P_1) \geq U(f,Q) \geq L(f,Q) \geq L(f,P_2)$$

Thus, every upper sum for a fixed f dominates every lower sum.

Definition 1.2.2 — Riemann Integrable. If $\inf\{U(f,P) : \forall P\} = \sup\{L(f,P) : \forall P\}$, we say that f is Riemann integrable over $[a,b]$ and write

$$R-\int_a^b f = \inf_P U(f,P)$$

R

1. The notation is to emphasize it is Riemann integral
2. Leads to FTC, continuous function (minus a measure zero set) integrability
3. Later a characterization will be given

Definition 1.2.3 — Characteristic Function of A. Given $A \subseteq X$, denote $\chi_A : X \to \{0,1\}$ by

$$\chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

(This is pretty much indicator function in statistics.)

■ **Example 1.5 — Dirichlet Function-Not R Integrable.** Let $f = \chi_{\mathbb{Q}}$ and $[a,b]=[0,1]$. Let P be any partition of $[0,1]$, we always have $\sup f\big|_{[x_i,x_{i+1}]} = 1$ while $\inf f\big|_{[x_i,x_{i+1}]} = 0$. Thus, $U(f,P)=1$ and $L(f,P)=0$ for all P. f is not Riemann integrable. ■

■ **Example 1.6 — Poor Behaviour Under Limits I-Not R Integrable.** Consider the sequence of functions $\{g_n\}$ on $[0,1]$ such that $g_n(0)=0, g_n\left(\frac{1}{2n}\right)=n$ and $g_n(x)=0$ for $x \in \left[\frac{1}{n},1\right]$ and g_n is linear on $\left[0,\frac{1}{2n}\right]$ and $\left[\frac{1}{2n},\frac{1}{n}\right]$. The picture is given below. Note that each g_n is clearly continuous on $[0,1]$

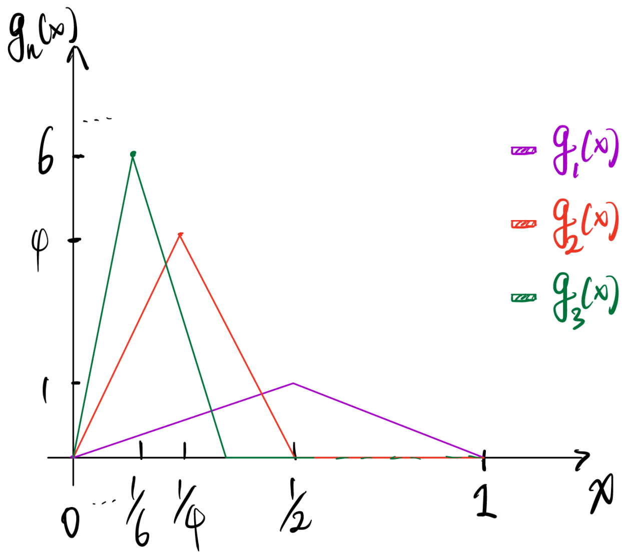


Figure 1.2.1: $g_n(x)$

and converging to 0 pointwise. Using the common area interpretation, we have $R-\int_0^1 g_n = \frac{1}{2}$ for all n. Thus, $\lim_{n\to\infty} R-\int g_n = \frac{1}{2}$. However, $R-\int_0^1 \lim_{n\to\infty} g_n = 0$. **We cannot switch the limit and integral without uniform convergence.** ■

R We know that Riemann integral itself is a limiting process, whenever we are switching two limiting process, there should be some concerns.

■ **Example 1.7 — Poor Behaviour Under Limits II-Not R Integrable.** Let $\{r_n\}_{n=1}^{\infty} = \mathbb{Q} \cap [0,1]$ (this is doable since $\mathbb{Q}$ is countable). Define

$$f_n(x) = \begin{cases} 1 & x = r_j, j \leq n \\ 0 & \text{otherwise} \end{cases}$$

As f_n has only countably many discontinuities, it is Riemann integrable and the integral evaluates to 0. However, $\lim_{n\to\infty} f_n = \chi_{\mathbb{Q}\cap[0,1]}$ which is not even Riemann integrable this time. ■

> **R** From these two examples above, we know that
>
> 1. Interchange limit and Riemann integral might result in different values
> 2. Interchange limit and Riemann integral might also just make things not defined

1.2.1 Intuition for Lebesgue Integral

In the notion of Lebesgue Integral, Example 1.7 will have an integral value of 0 and it can cope with pointwise limits of Lebesgue integrable functions.

Lebesgue's Idea:

Instead of partitioning the domain, we partition of the range of the function and look at the preimage of these little sub-intervals.

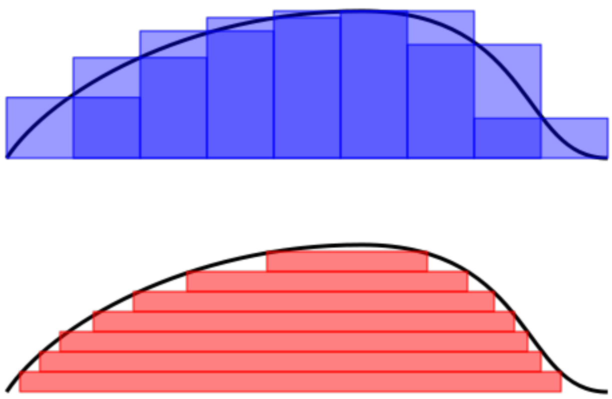


Figure 1.2.2: *Riemann Integral (Top) vs. Lebesgue Integral (Bottom)*

Definition 1.2.4 — Preimage. For $\operatorname{ran}(f) = [A,B]$, consider teh partition $A = y_1 < \cdots < y_N = B$, then

$$f^{-1}(S) = \{x \in \operatorname{dom}(f) : f(x) \in S\}$$

is the preimage of S under the function f. We say

$$E_i = f^{-1}(y_{i-1}, y_i]$$

■ **Example 1.8** For $f = \chi_{\mathbb{Q}}$, the preimage sets are

$$E = \begin{cases} \mathbb{Q} & \{1\} \subseteq S, \{0\} \not\subseteq S \\ \mathbb{Q}^c & \{0\} \subseteq S, \{1\} \not\subseteq S \\ \mathbb{R} & \{0,1\} \subseteq S \\ \emptyset & \{0,1\} \not\subseteq S \end{cases}$$

■

In general, when we are taking the partition of the range small enough, we can have the preimage sets to be a partition of the domain of f, then

$$f \sim \sum y_i \chi_{E_i}$$

(the $\sim$ notation means approximation) then, the question comes to how to rigorously define and evaluate the following quantity.

$$\lim \sum_i y_i \text{“length of”} E_i$$

1.3 Ideal Notion of a Measure Function

To extend the notion of length of intervals to much general sets. We want to propose a function m

$$m : \mathscr{P}(\mathbb{R}) \to [0,\infty] = \mathbb{R}^+ \cap \{\infty\}$$

with the following properties:

1. $m(\text{interval}) =$interval length, and $m(\emptyset) = 0$
2. $\sigma-$**additivity:** Let $\{E_n\}_{n=1}^{\infty}$ be disjoint sets, then

$$m\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} m(E_n)$$

> R This will be important for limit theorems.

3. **Translation Invariance:** for $E \subseteq \mathbb{R}$ and $y \in \mathbb{R}$, let $E + y = \{e + y : e \in E\}$, then $m(E) = m(E + y)$

> R Sliding the set along the real line does not change the length

Corollary 1.3.1 — Monotonicity from Property 2. If $A \subseteq B$, then $m(A) \leq m(B)$.

Proof. Since $B = A \dot{\cup} (B \backslash A)$, by $\sigma-$additivity, we have

$$m(B) = m(A) + m(B \backslash A) \geq m(A)$$

■

We will show that such an m function does not exist due to AOC.

Exercise 1.3 Define a relation on $\mathbb{R}$ by the rule that $x \sim y$ if $x - y \in \mathbb{Q}$. Verify that this is a equivalence relationship.

Solution:

1. **Reflexive:** for any $x \in \mathbb{Q}$, we clearly have $x \sim x$ since $x - x = 0 \in \mathbb{Q}$
2. **Symmetric:** for any $x, y \in \mathbb{Q}$ such that $x \sim y$ then $x - y \in \mathbb{Q}$, then $y - x = -(x - y) \in \mathbb{Q}$.

Thus, $y \sim x$

3. **Transitive:** given $x \sim y, y \sim z$, then

$$\underbrace{x-y}_{\in\mathbb{Q}}+\underbrace{y-z}_{\in\mathbb{Q}}=x-z\in\mathbb{Q}$$

thus, $x \sim z$. ■

Theorem 5 — Such m cannot exist. *Proof.* Using the equivalence relation on $\mathbb{R}$, we have a partition of $\mathbb{R}$ with equivalent classes $x_0+\mathbb{Q}, x_0 \in \mathbb{R}$.
Since $\overline{\mathbb{Q}}=\mathbb{R}$, we know each $x_0+\mathbb{Q}\cap\left[0,\frac{1}{2}\right]\neq\emptyset$. By AOC, we can pick out an element out of each intersection and denote the collection as E. For $x_1 \neq x_2 \in \mathbb{Q}$, we consider $E+x_1, E+x_2$ two sets

Claim: $E+x_1$ and $E+x_2$ are disjoint
Proof. Suppose note, then there exists $e_1, e_2 \in E$ such that $e_1+x_1=e_2+x_2 \implies e_1-e_2=x_2-x_1\in\mathbb{Q}\implies e_1\sim e_2$. But $e_1\neq e_2$ is not possible since they are picked from disjoint equivalence classes. Thus, $e_1=e_2 \implies x_1=x_2$, again, this yields a contradiction. ■

Then, by $\sigma-$additivity and translation invariance, we have

$$m\left(\bigcup_{x\in\mathbb{Q}}E+x\right)\underset{\sigma-\text{additivity}}{=}\sum_{x\in\mathbb{Q}}m(E+x)\underset{\text{Tranlation Invariance}}{=}\sum_{x\in\mathbb{Q}}m(E)=0\text{ or }\infty$$

It will be 0 if $m(E)=0$, otherwise, it will diverge to ∞.

Claim: $m(E)>0$
Proof. Note that $\cup_{x\in\mathbb{Q}}E+x=\mathbb{R}$ since any $y\in\mathbb{R}$, we know $y\sim e\in E$, then, $y=e+x$ for some $x\in\mathbb{Q}$. Thus, $y\in E+x$. Then,

$$m\left(\bigcup_{x\in\mathbb{Q}}E+x\right)=m(\mathbb{R})=\infty$$

this implies $m(E)>0$. ■

Due to countability, we can let $\mathbb{Q}\cap\left[0,\frac{1}{2}\right]=\{r_i\}_{i=1}^{\infty}$ be some sequence. Then, by $\sigma-$additivity and translation invariance, we have

$$m\left(\bigcup_{i=1}^{\infty}E+r_i\right)=\sum_{i=1}^{\infty}m(E+r_i)=\sum_{i=1}^{\infty}m(E)=\infty$$

but by construction, we know that $E\subseteq\left[0,\frac{1}{2}\right]\subseteq[0,1]$, so $E+r_i\subseteq[0,1]$. Then, by monotonicity, we have

$$m\left(\bigcup_{i=1}^{\infty}E+r_i\right)\leq m([0,1])=1$$

This yields a contradiction. ■

R **What Went Wrong?**

1. We are not giving up on translation invariance
2. We are not giving up on the non-negative measure range
3. We are not giving up on the monotonicity
4. We are giving up on $\sigma-$additivity partially by letting m defined on a subset of the power set—$\sigma-$algebra. Even if we just do finite additivity, we still cannot make m work (Good question, but hard proof)

We are going to define a measure function on $\mathscr{P}(\mathbb{R})$ without $\sigma-$additivity—**Lebesgue Outer Measure** (this will give us σ-subadditivity, but not good enough for us), then we will restrict its domain to Lebesgue measurable sets to retrieve the $\sigma-$additivity.

Exercise 1.4 Show m does not exist even if we just have finite $\sigma-$additivity.
Solution: According to *Real Analysis* by Halsey Royden, there was a proof by contradiction on Page 49 for this. ■

1.4 Construction of Lebesgue Outer Measure

The function that was discussed in the last part of the last section, m^*, is called **outer Lebesgue measure**.

Definition 1.4.1 — Outer Lebesgue Measure. Given $A \subseteq \mathbb{R}$, let

$$\mathscr{C}(A) := \{\{I_n\}_{n=1}^{\infty} \text{ where } I_n \text{ are open intervals with } A \subseteq \cup_{n=1}^{\infty} I_n\}$$

We write $l(I)$ for the length of an interval I. Define

$$m^*(A) = \inf\left\{\sum_{n=1}^{\infty} l(I_n) : \{I_n\}_{n=1}^{\infty} \in \mathscr{C}(A)\right\}$$

Does this notion of outer Lebesgue measure give us what it promised in the last section?

Exercise 1.5 — Easy Properties of m^*.

1. $m^* : \mathscr{P}(\mathbb{R}) \to [0,\infty]$, $m^*(\phi) = 0, m^*(\{x\}) = 0$

   Proof. (a) Since the length function is non-negative and $A \subseteq \mathbb{R}$, we know that $m^* : \mathscr{P}(\mathbb{R} \to [0,\infty])$.
   (b) If $A =$, the open cover can just be $\emptyset$ and $l(\emptyset) = 0$ implies that $m^*(A) = 0$ by taking the infimum. As infimum, the smallest it can be and it has been is 0.
   (c) If A is a singleton set, we can consider the interval $(x-\varepsilon, x+\varepsilon)$ which is an open cover of the discrete point for any $\varepsilon > 0$. Then,

   $$m^*(A) = \inf\left\{\sum_{n=1}^{\infty} l(I_n) : \{I_n\}_{n=1}^{\infty} \in \mathscr{C}(A)\right\} \le \inf\{l((x-\varepsilon, x+\varepsilon)), \forall \varepsilon > 0\} = 0$$

   Thus, $m^*(A) = 0$. ■

2. **Monotonicity:** $m^*(A) \le m^*(B)$ if $A \subseteq B$

   Proof. Since $A \subseteq B$, every countable open cover of B is a countable open cover of A, so

$\mathscr{C}(B) \subseteq \mathscr{C}(A)$. Thus, by infimum property, we have

$$m^*(A) = \inf\left\{\sum_{n=1}^{\infty} l(I_n) : \{I_n\}_{n=1}^{\infty} \in \mathscr{C}(A)\right\} \leq \inf\left\{\sum_{n=1}^{\infty} l(I_n) : \{I_n\}_{n=1}^{\infty} \in \mathscr{C}(B)\right\} = m^*(B)$$

■

3. **Translation Invariance:** m^* is translation invariant

Proof. This is clear from the definition since l itself is translation invariant. ■

Proposition 1.4.1 $m^*(I) = l(I)$ if I is an interval.

Proof. We shall prove the case when $I = [a,b]$. For other types of interval, it is clear enough from the easy properties in the previous exercise. We can build other results upon this special case by using monotonicity.
Let $\varepsilon > 0$, $(a-\varepsilon, b+\varepsilon)$ is an open cover of $[a,b]$. Then,

$$m^*(I) \leq l((a-\varepsilon, b+\varepsilon)) = b - a + 2\varepsilon$$

Since $\varepsilon > 0$ is arbitrary. We have $m^*(I) \leq b - a$.
For the other side, we will show that whenever $\{I_n\}$ is an open cover of I, then $\sum_{n=1}^{\infty} l(I_n) \geq b - a$. Since $\mathbb{R}$ is metric space and I is a closed and bounded set, therefore compact. There exists a finite subcover $I_1, \ldots, I_N$ such that

$$I \subseteq \sum_{n=1}^{N} I_n$$

then, we also have

$$\sum_{n=1}^{\infty} l(I_n) \geq \sum_{n=1}^{N} l(I_n)$$

it suffices to show that $\sum_{n=1}^{N} l(I_n) \geq b - a$ whenever $I \subseteq \sum_{n=1}^{N} I_n$. Say $a \in I_{n_1} = (a_1, b_1)$, if $b_1 > b$, we have

$$\sum_{n=1}^{N} l(I_n) \geq l(I_{n_1}) = b_1 - a_1 \geq b - a$$

and we are done. Suppose not, then, we can sequentially look at the next I_{n_2}, to see if b is contained. Since there are only finitely many of them, we will find $b \in I_{n_k}$ for some k. Then,

$$\sum_{n=1}^{N} l(I_n) \geq \sum_{j=1}^{k} l(I_{n_j}) \geq \sum_{j=1}^{k} b_j - a_j \geq b - a$$

This process can be visualized as below.

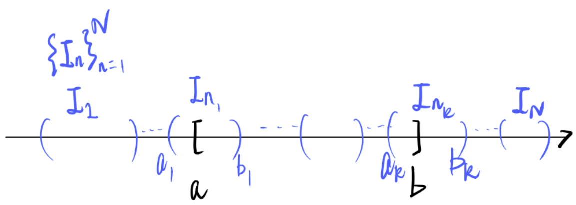


Figure 1.4.1: *The process to find* I_{n_k}

Thus, we have $m^*(I) = b - a$ as desired. ■

> **R** We are reminded that so far m^* can only perform $\sigma-$sub-additivity. What do we mean by that?

Proposition 1.4.2 — $\sigma-$sub-additivity. For all $A_k \subseteq \mathbb{R}$ (not necessarily disjoint), we have

$$m^*\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} m^*(A_k)$$

Proof. For $\varepsilon > 0$, for each A_k, we choose $\{I_{n,k}\}_{n=1}^{\infty} \in \mathscr{C}(A_k)$ such that $\cup_{n=1}^{\infty} I_{n,k} \supseteq A_k$ and

$$\sum_{n=1}^{\infty} l(I_{n,k}) \leq m^*(A_k) + \varepsilon 2^{-k}$$

(This can be done by the infimum definition) Note that $\{I_{n,k}\}_{k,n=1}^{\infty}$ is a countable collection of open intervals whose union contains A and hence in $\mathscr{C}(A)$. Then,

$$m^*\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k,n=1}^{\infty} l(I_{n,k}) \leq \sum_{k=1}^{\infty} \left(m^*(A_k) + \varepsilon 2^{-k}\right) = \sum_{k=1}^{\infty} (m^*(A_k)) + \varepsilon$$

since $\varepsilon > 0$ is arbitrary, we are done. ■

Corollary 1.4.3 If $A = \{x_n\}_{n=1}^{\infty}$ is any countable set, then $m^*(A) \leq \sum_{n=1}^{\infty} m^*(\{x_n\}) = 0$.

Corollary 1.4.4 — Every set is close to be "open". For any $A \subseteq \mathbb{R}$ and for any $\varepsilon > 0$ there is an open set $O \supseteq A$ such that $m^*(O) \leq m^*(A) + \varepsilon$.

Proof. Choose open intervals $\{I_n\}_{n=1}^{\infty} \in \mathscr{C}(A)$ such that $\sum_{n=1}^{\infty} l(I_n) \leq m^*(A) + \varepsilon$ and put $O = \cup_{n=1}^{\infty} I_n$ which is open. Then, by $\sigma-$sub-additivity,

$$m^*(O) = m^*\left(\bigcup_{n=1}^{\infty} I_n\right) \leq \sum_{n=1}^{\infty} m^*(I_n) \underset{\text{it's an interval}}{=} \sum_{n=1}^{\infty} l(I_n) \leq m^*(A) + \varepsilon$$

■

1.5 Construction of Lebesgue Measure

1.5.1 Lebesgue Measurable Sets and Lebesgue Measure

Clearly, we cannot let m^* be on $\mathscr{P}(A)$, we need a special subset of it.

Definition 1.5.1 — Lebesgue Measurable. We say that $A \subseteq \mathbb{R}$ is Lebesgue measurable if for every $E \subseteq \mathbb{R}$ we have

$$m^*(E) = m^*\left(E \bigcap A\right) + m^*\left(E \bigcap A^c\right)$$

> **R** Also known as Caratheodory definition of Lebesgue measurability. Note that by sub-additivity, one direction is clear
>
> $$m^*(E) \leq m^*\left(E \bigcap A\right) + m^*\left(E \bigcap A^c\right)$$
>
> we only need to check the other direction.

■ **Example 1.9**
1. $A = \mathbb{R}$ or $\emptyset$
2. A^c if A is Lebesgue measurable since the definition is symmetric
3. Any set A with $m^*(A) = 0$. Then, any subset $E \subseteq A$ will have $m^*(E \cap A) \leq m^*(A) = 0$. Then,

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

■

Proposition 1.5.1 The interval (a, ∞) is Lebesgue measurable.

Proof. We need to show that for any $E \subseteq \mathbb{R}$, we have

$$m^*(E) = M^*(E \cap (a, \infty)) + m^*(E \cap (-\infty, a])$$

Let $\varepsilon > 0$ and let $\{I_n\} \in \mathscr{C}(E)$ such that

$$\sum_n l(I_n) \leq m^*(E) + \varepsilon$$

Let $I_n^+ = I_n \cap (a, \infty)$ and $I_n^- = I_n \cap (-\infty, a]$. Note that I_n^+ and I_n^- are either intervals or empty. Then,

$$l(I_n) = l(I_n^+) + l(I_n^-) = m^*(I_n^+) + m^*(I_n^-)$$

since $E \cap (a, \infty) \subseteq \cup_n I_n \cap (a, \infty) = \bigcup_n I_n^+$, by monotonicity and σ-sub-additivity,

$$m^*(E \cap (a, \infty)) \leq m^*\left(\bigcup_n I_n^+\right) \leq \sum_n m^*(I_n^+)$$

similarly, $E \cap (-\infty, a] \subseteq \cup_n I_n^+$ and

$$m^*(E \cap (-\infty, a]) \leq \sum_n m^*(I_n^-)$$

finally,

$$m^*(E \cap (a, \infty)) + m^*(E \cap (-\infty, a]) \leq \sum_n m^*(I_n^+) + m^*(I_n^-) = \sum_n l(I_n) \leq m^*(E) + \varepsilon$$

since $\varepsilon > 0$ arbitrarily, we have the desired result. ■

The class of all Lebesgue measurable sets have a structure called $\sigma-$algebra.

Definition 1.5.2 — σ-algebra. A collection Ω of subsets of $\mathbb{R}$ is called a $\sigma-$algebra if
1. $\emptyset \in \Omega$
2. **Closed under countable unions:** If $\{A_n\}_{n \in \mathbb{N}} \subseteq \Omega$, then $\cup_{n=1}^\infty A_n \in \Omega$
3. **Closed under complements:** If $A \in \Omega$, then $A^c \in \Omega$

■ **Example 1.10**
1. Smallest σ-algebra is $\{\emptyset, \mathbb{R}\}$
2. Largest $\sigma-$algebra is $\mathscr{P}(\mathbb{R})$

■

R

Closed under countable intersection:

Note that if $\{A_n\}_{n \in \mathbb{N}} \subseteq \Omega$, then $A_n^c \in \Omega, \forall n \in \mathbb{N}$. Then, $\cup_{n=1}^\infty A_n^c \in \Omega$ by def 2. Then, by De Morgan's Law, we have $\cup_{n=1}^\infty A_n^c = (\cap_{n=1}^\infty A_n)^c$. Thus, Ω is closed under countable intersections.

Exercise 1.6 Any intersection of $\sigma-$algebra is again a $\sigma-$algebra.
Solution:
It suffices to show for two $\sigma-$algebras, Ω_1 and Ω_2. Consider $\Omega = \Omega_1 \cap \Omega_2$

1. Since $\emptyset \in \Omega_1, \Omega_2$, certainly $\emptyset \in \Omega$
2. Let $\{A_n\} \subseteq \Omega$, then $\cup_n A_n \in \Omega_1$ and Ω_2 since $A_n \in \Omega_1, \Omega_2, \forall n \in \mathbb{N}$. Thus, $\cup_n A_n \in \Omega$.
3. Let $A \in \Omega$, then $A^c \in \Omega_1, \Omega_2$, thus, $A^c \in \Omega$.

■

For the real line, we particularly cares more about the following.

Definition 1.5.3 — Borel $\sigma-$algebra.
1. The Borel $\sigma-$algebra is the intersection of all the $\sigma-$algebras containing all the open sets in $\mathbb{R}$
2. A Borel set is any set in the Borel $\sigma-$algebra.

R Borel $\sigma-$algebra is by definition the smallest $\sigma-$algebra that contains all the open sets in $\mathbb{R}$. This is supported by the fact that every open set in $\mathbb{R}$ is a countable union of open intervals (This was proved in a PMATH351 assignment offered by Prof.Marcoux, the Chichen Mathematician). We also say the Borel $\sigma-$algebra is the smallest $\sigma - algebra$ generated by the open sets/open intervals. We can explicitly generate $\sigma-$algebra based on some given elements.

All the G_δ, F_σ sets are in the Borel $\sigma-$algebra.

Exercise 1.7
1. Show that the Borel $\sigma-$algebra is the smallest $\sigma-$algebra generated by the intervals of the form (a, ∞)
Solution:
(a) We first show that the Borel $\sigma-$algebra can be generated by the intervals of (a, ∞). We know that (a, ∞) is Lebesgue measurable for all $a \in \mathbb{R}$. We just need to show (a, ∞) can generate all open intervals. In particular, $(a,b) = (a, \infty) \cap \left(\cup_n \left(b - \frac{1}{n}, \infty\right)^c\right)$.
(b) Its the intersection of every $\sigma-$algebra, for sure, it is the smallest.
2. Show that every interval is a Borel set.
Solution: Since $(-\infty, a), (a,b), (a, \infty)$ are open, they must be Borel sets. It suffices for us to show that
$$[a,b] = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, b + \frac{1}{n}\right)$$
$$(a,b] = \bigcap_{n=1}^{\infty} \left(a, b + \frac{1}{n}\right)$$
$$[a,b) = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, b\right)$$
which are set equivalence exercises.
3. Find the cardinality of the Borel $\sigma-$algebra?
Solution: I don't know yet.

■

Finally, we can retrieve the $\sigma-$additivity.

Theorem 6 — Caratheodory Extension Theorem. The set of Lebesgue measurable sets, denoted by $\mathscr{M}$, is a $\sigma-$algebra which contains the Borel sets and includes all sets of outer

Lebesgue measure zero. Moreoever, if $A_n \in \mathscr{M}$ are disjoint, then

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

Thus, m^* restricted to $\mathscr{M}$ is $\sigma-$additive.

R

1. The hard task here is to show that countable union of Lebesgue measurable sets is still Lebesgue measurable.
2. We have (a,∞) Lebesgue measurable, together with the exercise, it is clear that $\mathscr{M}$ contains all the Borel sets.
3. The additive property is also left to prove.

Definition 1.5.4 — Lebesgue Measure in $\mathbb{R}$. By Lebesgue measure m on $\mathbb{R}$, we mean m^* restricted to $\mathscr{M}$.

■ **Example 1.11 — Cantor Set-Uncountable Set of Lebesgue Measure Zero.** The formulation of a Cantor set is intuitive by graph. It has been shown in MATH147/PMATH351 that the Cantor set is compact and therefore measurable. Its interior is empty and every point is an accumulation point, and it is uncountable (perfect set in $\mathbb{R}$ is always uncountable). Graphically, you can see that it will be Lebesgue measure zero. Or, if you believe in mathematics, since C is closed it is a Borel set, we have

$$m(C) \le m(C_n) = 2^n 3^{-n}, \forall n \Longrightarrow m(C) = 0$$

even though its cardinality is as large as the real line but the measure is as small as the empty.

Figure 1.5.1: *Cantor Set*

■

We summarize some of the useful properties of Lebesgue Measure as below (a lot of them are inherited from outer Lebesgue measure)

Theorem 7 — Properties of Lebesgue Measure.
1. $m(I) =$length of the interval I
2. If $A \in \mathscr{M}$, then $A+t \in \mathscr{M}$ for all $t \in \mathbb{R}$ and $m(A) = m(A+t)$
3. m is $\sigma-$additive (need to be disjoint)
4. m is $\sigma-$sub-additive (not necessarily disjoint)
5. If $A, B \in \mathscr{M}$ and $A \subseteq B$, then $A^c \cap B \in \mathscr{M}$ and since $B = A \cup (A^c \cap B)$, we have

$$m(B) = m(A) + m(A^c \cap B) = m(A) + m(B \backslash A)$$

if $m(A) < \infty$, then $m(B \backslash A) = m(B) - m(A)$
6. m is monotonic

7. A set E has Lebesgue measure zero if and only if for every $\varepsilon > 0$ there are open intervals I_n such that $\cup_n I_n \supseteq E$ and $\sum_n l(I_n) < \varepsilon$
8. $m(C) < \infty$ for any compact measurable set C. Since it can be covered by $[-N,N]$ for large enough N.

Exercise 1.8 Show that is E is measurable and $\varepsilon > 0$, then there is an open set $O \supseteq E$ such that $m(O \backslash E) < \varepsilon$. (Hint. do this first with the assumption that $m(E) < \infty$)

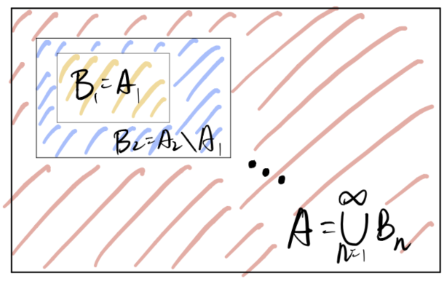
Solution:

Proof.

1. If $m(E) < \infty$, by a corollary in class, we know that there exists an open set $O \supseteq E$ such that
$$m^*(O) \le m^*(E) + \varepsilon$$
since every open set is in the Borel $\sigma-$algebra, we actually have
$$m(O) \le m(E) + \varepsilon$$
since $m(E) < \infty$, we have $m(O \backslash E) = m(O) - m(E) \le \varepsilon$
2. If $m(E) = \infty$, for each $n \ge 1$, define $E_n = E \cap [-n,n]$ where $[-n,n]$ is in the Borel σ algebra, thus measurable. Then, $m(E_n) \le 2n < \infty$. Then, by the first case, we have $O_n \supseteq E_n$ as open set for each $n \ge 1$ such that $m(O_n \backslash E_n) \le \frac{\varepsilon}{2^n}$. We let $O = \cup_{n=1}^{\infty} O_n$ be the countable union of open set (still open), then
$$E = \bigcup_{n=1}^{\infty} E_n \subseteq \bigcup_{n=1}^{\infty} O_n = O$$
consider
$$O \backslash E = \bigcup_{n=1}^{\infty} (O_n \backslash E) = \bigcup_{n=1}^{\infty} \left(O_n \backslash \bigcup_{i=1}^{\infty} E_i \right) \subseteq \bigcup_{n=1}^{\infty} (O_n \backslash E_n)$$
Then, by σ-sub-additivity and monotonicity, we have
$$m(O \backslash E) \le \sum_{n=1}^{\infty} m(O_n \backslash E_n) \le \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon$$

This completes the proof of the statement. ■

Proposition 1.5.2 — Continuity of Measure. If $A_1 \subseteq A_2 \subseteq \cdots \subseteq \cup_{n=1}^{\infty} A_n = A \in \mathscr{M}$ for all n, then $A \in \mathscr{M}$ and
$$m(A) = \lim_n m(A_n)$$



Proof. Let $B_1 = A_1$ and for $n > 1$, let $B_n = A_n \backslash A_{n-1}$. Then, for each $B_n \in \mathscr{M}$ and these sets are disjoint. Furthermore, $\cup_{n=1}^{\infty} B_n = A$ and $\cup_{k=1}^{n} B_k = A_n$, then by $\sigma-$additivity,

$$m(A_n) = \sum_{k=1}^{n} m(B_k) \Longrightarrow \sum_{k=1}^{\infty} m(B_k) = m(A)$$

■

Proposition 1.5.3 — Downward Continuity of Measure. If $A_1 \supseteq A_2 \supseteq \cdots \supseteq \cap_{n=1}^{\infty} A_n = A$ and $A_n \in \mathscr{M}$ for all n, then $A \in \mathscr{M}$. If, in addition, $m(A_1) < \infty$, then

$$m(A) = \lim_n m(A_n)$$

Proof. Let $B_n = A_1 \backslash A_n$ and $B = A_1 \backslash A$, then $B_n \in \mathscr{B}$, $B_n \subseteq B_{n+1}$ and $\cup_{n=1}^{\infty} B_n = B$. By the first part, $m(B) = \lim_n m(B_n)$. But also

$$m(B) = m(A_1) - m(A)$$
$$m(B_n) = m(A_1) - m(A_n)$$

■

■ **Example 1.12 — $m(A_1) < \infty$ is necessary.** Consider $A_n = (n, \infty)$. Clearly that $A_{n+1} \subseteq A_n$ and $\cap_{n=1}^{\infty} A_n = \emptyset$. Here $m(A_n) = \infty$ for all n, while $m(\cap_{n=1}^{\infty} A_n) = 0$. ■

Exercise 1.9 Show that for any $E \in \mathscr{M}$.

$$m(E) = \sup\{m(K) : K \subseteq E, K \text{ is compact}\}$$

Solution: ■

R What we have learned in class about Lebesgue measurability is actually the Caratheodory's version of it. The original Lebesgue's definition involves the inner Lebesgue measure that is defined to be

$$m_*(E) = \sup\{m(K) : K \subseteq E, K \text{ is compact}\}$$

If a set E is Lebesgue measurable in the Lebesgue sense, then

$$m^*(E) = m_*(E)$$

The exercise above proved that these two senses (Caratheodory and Lebesgue) of measurable sets are indeed equivalent.

1.5.2 Lebesgue's Characterization of Riemann Integrability

The following theorem really shows that Riemann integrability is still viable under the Lebesgue realm.

Theorem 8 Let $f : [a,b] \to \mathbb{R}$ be a bounded function. Then f is Riemann integrable over $[a,b]$ if and only if $m(E) = 0$ where E is the set of discontinuities of f over $[a,b]$.

Proof. (a) Assume f is Riemann integrable over $[a,b]$. Let $k \in \mathbb{N}$ and let

$$E_k = \left\{ z \in [a,b] : \forall \delta > 0, \exists x, y \in [z-\delta, z+\delta] \text{ s.t. } |f(x) - f(y)| > \frac{1}{k} \right\}$$

(i) We claim that $E = \cup_{k=1}^{\infty} E_k$

i. First, we show that $E \subseteq \cup_{k=1}^{\infty} E_k$. For $z \in E$, there exists $\varepsilon > 0$ such that $\forall \delta > 0$, there exists x with $|x-z|<\delta$ implies $|f(x)-f(z)|\geq\varepsilon$. We can pick $k\in\mathbb{N}$ such that $\varepsilon>\frac{1}{k}$. Thus, $z\in E_k$ and the union.

ii. Second, we show that $E\supseteq\cup_{k=1}^{\infty}E_k$, let $z\in\cup_{k=1}^{\infty}E_k$. In particular, $z\in E_k$ for k. For the sake contradiction, say f is continuous at z. We let $\varepsilon=\frac{1}{4k}>0$, then there exists $\delta>0$ such that $x\in(z-\delta,z+\delta)$ implies

$$|f(x)-f(z)|<\frac{1}{4k}$$

but $z\in E_k$, so for $\delta_1=\frac{\delta}{4}>0$, there exists $c,d\in[z-\delta_1,z+\delta]\subseteq(z-\delta,z+\delta)$ such that

$$|f(c)-f(d)|>\frac{1}{k}$$

also, we have

$$|f(c)-f(z)|<\frac{1}{4k},\qquad |f(d)-f(z)|<\frac{1}{4k}$$

but then, by triangle inequality,

$$\frac{1}{2k}>|f(c)-f(z)|+|f(z)-f(d)|\geq|f(c)-f(d)|>\frac{1}{k}$$

this yields a contradiction. Thus, $z\in E$. This sums up to $E\supseteq\cup_{k=1}^{\infty}E_k$

Thus,

$$E=\cup_{k=1}^{\infty}E_k$$

(ii) Fix $\varepsilon>0$ and $k\in\mathbb{N}$. Let $P:a=x_0<x_1<\cdots<x_n=b$ we a partition of $[a,b]$ such that

$$U(f,P)-L(f,P)<\frac{\varepsilon}{k}$$

(this exists since f is Riemann integrable over $[a,b]$). Let $J_k=\{i:(x_{i-1},x_i)\cap E_K\neq\emptyset\}$

Claim: $\sum_{i\in J_k}x_i-x_{i-1}\leq\varepsilon$

Proof. Note that we can write

$$\sum_{i\in J_k}\left(\sup_{x\in[x_{i-1},x_i]}\{f(x)\}-\inf_{x\in[x_{i-1},x_i]}\{f(x)\}\right)(x_i-x_{i-1})\leq U(f,P)-L(f,P)<\frac{\varepsilon}{k}$$

For each $i\in J_k$, we can find $z\in E_k\cap(x_{i-1},x_i)\subseteq[x_{i-1},x_i]$ such that for δ small enough to have $c_i,d_i\in[x_{i-1},x_i]$ and $c_i,d_i\in[z-\delta,z-\delta]$ such that $|f(c_i)-f(d_i)|>\frac{1}{k}$. Then,

$$\frac{1}{k}<|f(c_i)-f(d_i)|\leq\sup_{x\in[x_{i-1},x_i]}\{f(x)\}-\inf_{x\in[x_{i-1},x_i]}\{f(x)\},\forall i\in J_k$$

then,

$$\sum_{i\in J_k}\left(\frac{1}{k}\right)(x_i-x_{i-1})<\sum_{i\in J_k}|f(c_i)-f(d_i)|(x_i-x_{i-1})$$

$$\leq\sum_{i\in J_k}\left(\sup_{x\in[x_{i-1},x_i]}\{f(x)\}-\inf_{x\in[x_{i-1},x_i]}\{f(x)\}\right)(x_i-x_{i-1})<\frac{\varepsilon}{k}$$

Thus,

$$\sum_{i\in J_k}(x_i - x_{i-1}) \le \varepsilon$$

■

Then, $E_k = \cup_{i\in J_k}(x_{i-1},x_i)\cap E_k \subseteq \cup_{i\in J_k}(x_{i-1},x_i)$, by mononicity and $\sigma-$sub-additivity of the outer Lebesgue measure, we have

$$m^*(E_k) \le m^*(\cup_{i\in J_k}(x_{i-1},x_i)) \le \sum_{i\in J_k} m^*((x_{i-1},x_i)) = \sum_{i\in J_k}(x_i - x_{i-1}) \le \varepsilon$$

thus, $m^*(E_k) \le \varepsilon$. Since $\varepsilon > 0$ is arbitrary, $m^*(E_k) = 0$, then by Example 3 in class, you know E_k is Lebesgue measurable. Then, note that

$$\bigcup_{k=1}^{N} E_k = E_N$$

since for $z \in E_k$, for any $\delta > 0$, there exists $x,y \in [z-\delta, z+\delta]$ such that $|f(x) - f(y)| > \frac{1}{k} > \frac{1}{k+1}$. Thus, $E_k \subseteq E_{k+1}$. Inductively, we have the result. Then, by continuity of measure, E is Lebesgue measurable and

$$m(E) = \lim_k m(E_k) = \lim_k 0 = 0$$

as desired.

(b) Conversely, assume $m(E) = 0$. Fix $\varepsilon > 0$ and obtain open intervals $\{I_n\}_{n=1}^{\infty}$ such that $E \subseteq \cup_n I_n$ and $\sum_n l(I_n) < \frac{\varepsilon}{4C}$ where $|f(x)| \le C$ for all x (we can find these by the infimum property and f is bounded).

(i) Let $\eta > 0$. Since $E \subseteq \cup_n I_n$, we know that f is continuous on T. Thus, for every $z \in T$, there exists $\delta_z > 0$ such that $|z-y| < \delta_z$ implies $|f(z) - f(y)| < \eta$. Also, let $Q_z := \{y \in [a,b] : |z-y| < \delta_z\}$, which is an open set. Then,

$$T \subseteq \bigcup_{z\in T} Q_z$$

since T is compact, there exists a finite subcover, $Q_{z_1}, \cdots, Q_{z_N}$ such that

$$T \subseteq \bigcup_{i=1}^{N} Q_{z_i}$$

then, we can pick $\delta = \min\{\delta_{z_1}, \ldots, \delta_{z_N}\}$ so that for $x \in T$ and $y \in [a,b]$ if $|x-y| < \delta$, we have

$$|f(x) - f(y)| < \eta$$

as desired.

(ii) Let $\eta = \frac{\varepsilon}{4(b-a)}$ and get $\delta > 0$ from (b)(i). Let $P : a = x_0 < x_1 < \cdots < x_n = b$ be a partition of $[a,b]$ with subintervals of length $< \delta$.

Claim: $U(f,P)-L(f,P)<\varepsilon$

Proof. Let $H=\{i:[x_{i-1},x_i]\cap T\neq\emptyset\}$. Then, within these subintervals, we have the property that was proved in (b)(i). And let $B=\{0,\cdots,n\}\backslash H$. Then,

$$\begin{aligned}
U(f,P)-L(f,P) &= \sum_{i\in H}\left(\sup_{x\in[x_{i-1},x_i]}\{f(x)\}-\inf_{x\in[x_{i-1},x_i]}\{f(x)\}\right)(x_i-x_{i-1})\\
&\quad+\sum_{i\in B}\left(\sup_{x\in[x_{i-1},x_i]}\{f(x)\}-\inf_{x\in[x_{i-1},x_i]}\{f(x)\}\right)(x_i-x_{i-1})\\
&\leq \frac{\varepsilon}{4(b-a)}\sum_{i=1}^{n}(x_i-x_{i-1})+2C\sum_{i\in B}(x_i-x_{i-1})\\
&\leq \frac{\varepsilon}{4}+2C\sum_{n}l(I_n)\\
&< \frac{\varepsilon}{4}+\frac{\varepsilon}{2}=\frac{3\varepsilon}{4}<\varepsilon
\end{aligned}$$

We are done after the following two holes that we need to fill.

Claim: $\sup_{x\in[x_{i-1},x_i]}\{f(x)\}-\inf_{x\in[x_{i-1},x_i]}\{f(x)\}\leq\eta$ **for** $i\in H$

Proof. It is clear that for $i\in H$, $|f(x)-f(y)|<\eta$ for any $x,y\in[x_{i-1},x_i]$. Thus,

$$f(x)<\eta+f(y),\forall x,y\in[x_{i-1},x_i]\implies\sup_{x\in[x_{i-1},x_i]}\{f(x)\}\leq\eta+f(y),\forall y\in[x_{i-1},x_i]$$

then,

$$\inf_{x\in[x_{i-1},x_i]}\{f(x)\}\geq\sup_{x\in[x_{i-1},x_i]}\{f(x)\}-\eta$$

rearrange, we have

$$\sup_{x\in[x_{i-1},x_i]}\{f(x)\}-\inf_{x\in[x_{i-1},x_i]}\{f(x)\}\leq\eta$$

■

Claim: $\sum_{i\in B}(x_i-x_{i-1})\leq\sum_n l(I_n)$

Proof. Note that $\cup_{i\in B}[x_{i-1},x_i]\subseteq\cup_n I_n$, then by σ-additivity of the disjoint (since they are subintervals of a partition and intervals are Lebesgue measurable), we have

$$\sum_{i\in B}(x_i-x_{i-1})=\sum_{i\in B}m^*([x_{i-1},x_i])=m^*\left(\bigcup_{i\in B}[x_{i-1},x_i]\right)$$

Then, by monotonicity and $\sigma-$sub-additivity,

$$m^*\left(\bigcup_{i\in B}[x_{i-1},x_i]\right)\leq m^*(\cup_n I_n)\leq\sum_n m^*(I_n)=\sum_n l(I_n)$$

the last equality is due to the fact that $\{I_n\}$ is a set of open intervals. ■

■

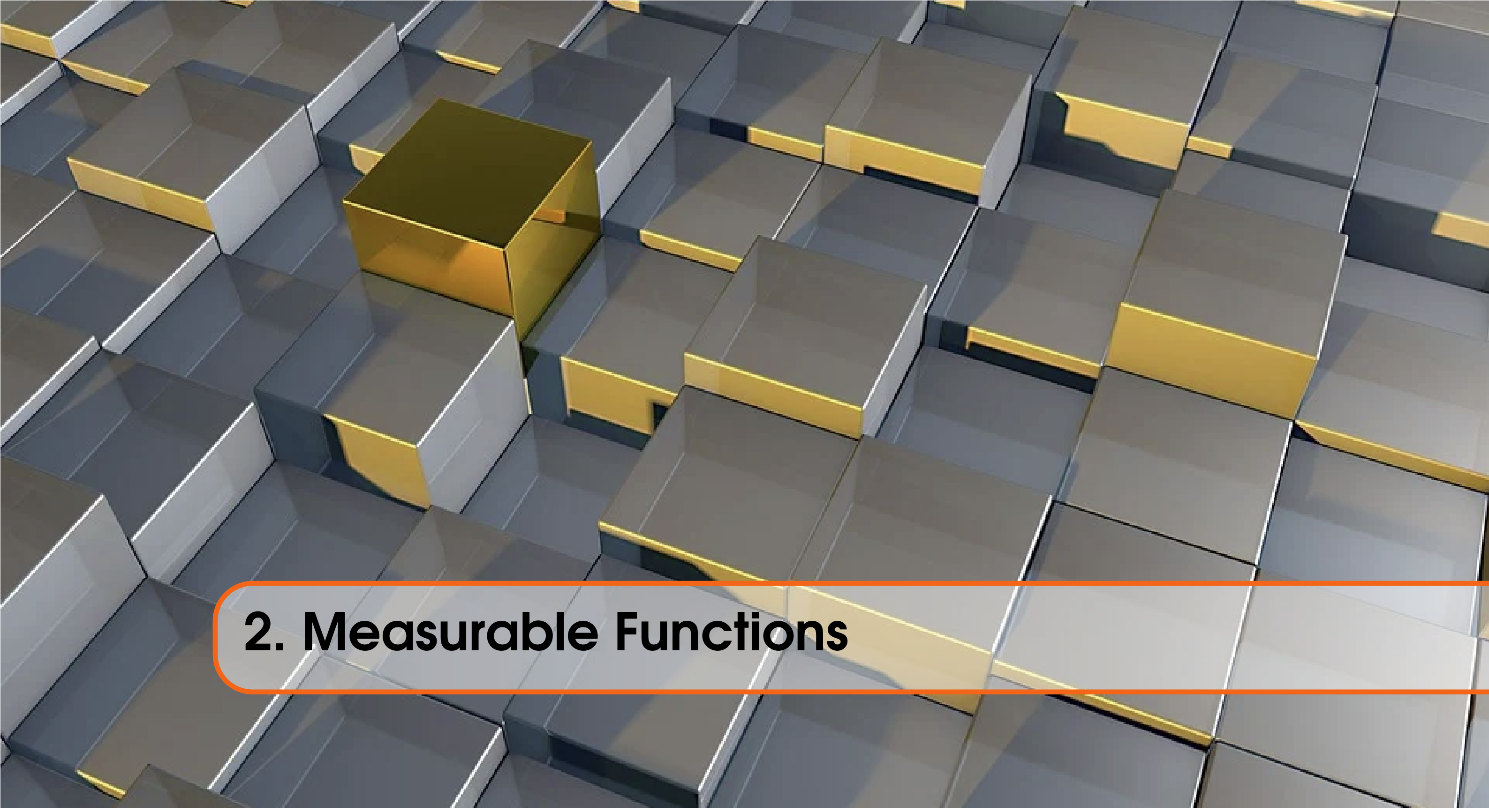
Thus, f is Riemann integrable.

■

■ **Example 1.13**

1. It is immediate from this that every function that is continuous, or has only finitely many discontinuities, or even has only countably many discontinuities, is Riemann integrable.
2. By the Increasing Function Theorem in MATH147, we know that it can only have countably many jump discontinuities, thus Riemann integrable
3. However, $f = \chi_{\mathbb{Q}}$ has uncountably many discontinuities, which is not Riemann integrable on any interval $[a, b]$.

■



2. Measurable Functions

2.1 Definitions

Definition 2.1.1 — Lebesgue Measurable Function. A function $f : X \to [-\infty, \infty]$ is called Lebesgue measurable if for every $\alpha \in \mathbb{R}$, the set

$$\{x \in X : f(x) < \alpha\} = f^{-1}([-\infty, \alpha))$$

is a Lebesgue measurable set.
A complex-valued function $f : X \to \mathbb{C}$ is called Lebesgue measurable if both its real and imaginary parts are Lebesgue measurable.

(R) Note that if f is real-valued, then $f^{-1}([-\infty, \alpha)) = f^{-1}((-\infty, \alpha))$

■ **Example 2.1**
1. Any constant function f is measurable since $f^{-1}([-\infty, \alpha))$ is either empty or all of $\mathbb{R}$ (when α is greater or equal to the value of the constant)
2. Any continuous, real-valued function f is measurable, since we have the remark above and the preimage of a continuous map on an open set is open, therefore Borel.

■

Exercise 2.1
1. Show that $\{x : f(x) < \alpha\}$ is measurable for all α if and only if $\{x : f(x) > \alpha\}$ is measurable for all $\alpha \in \mathbb{R}$

   Solution: We shall show that $\{x : h(x) < \alpha\}$ is Lebesgue measurable for all $\alpha \in \mathbb{R}$ if and only if $\{x : h(x) \leq \alpha\}$. Then, we will be done since

$$\{x : h(x) \leq \alpha\}^c = \{x : h(x) > \alpha\}$$

(a) Suppose $\{x : h(x) < \alpha\}$ for all $\alpha \in \mathbb{R}$. Then, for each $n \in \mathbb{N}$, we have

$$\left\{x : h(x) < \alpha + \frac{1}{n}\right\}$$

being Lebesgue measurable. Then, note that

$$\{x : h(x) \leq \alpha\} = \bigcap_{n=1}^{\infty} \left\{x : h(x) < \alpha + \frac{1}{n}\right\}$$

this is true since if $y \in \{x : h(x) \leq \alpha\}$, $h(y)$ is certainly less than any $\alpha + \frac{1}{n}$. Conversely, if $y \in \bigcap_{n=1}^{\infty} \left\{x : h(x) < \alpha + \frac{1}{n}\right\}$, $h(y) < \alpha + \frac{1}{n}$ for any $n \in \mathbb{N}$, which implies $h(y) \leq \inf\left\{\alpha + \frac{1}{n}\right\} = \alpha$, and we have $y \in \{x : h(x) \leq \alpha\}$.
Then, $\bigcap_{n=1}^{\infty} \left\{x : h(x) < \alpha + \frac{1}{n}\right\}$ is a countable intersection of measurable set, therefore, still measurable.

(b) Suppose $\{x : h(x) \leq \alpha\}$ is measurable for any $\alpha \in \mathbb{R}$. Then,

$$\left\{x : h(x) \leq \alpha - \frac{1}{n}\right\}$$

is Lebesgue measurable. And

$$\{x : h(x) < \alpha\} = \bigcup_{n=1}^{\infty} \left\{x : h(x) \leq \alpha - \frac{1}{n}\right\}$$

since if $y \in \{x : h(x) < \alpha\}$ such that $h(y) < \alpha$, we can always find a $n \in \mathbb{N}$ such that $h(y) \leq \alpha - \frac{1}{n} < \alpha$. Conversely, if y is in the union, then, $h(y) \leq \alpha - \frac{1}{n} < \alpha$ for any n. Thus, $y \in \{x : h(x) < \alpha\}$. Then, the right hand side is a countable union of measurable sets, so is the left hand side.

This concludes that $\{x : h(x) < \alpha\}$ is Lebesgue measurable for all $\alpha \in \mathbb{R}$ if and only if $\{x : h(x) \leq \alpha\}$ for all $\alpha \in \mathbb{R}$. Then, $\{x : h(x) < \alpha\}$ is Lebesgue measurable for all $\alpha \in \mathbb{R}$ if and only if $\{x : h(x) > \alpha\}$ for all $\alpha \in \mathbb{R}$.

2. Show that if f is measurable, then $f^{-1}(\{\infty\})$ is a measurable set.
Solution: This can be translated into

$$f^{-1}(\{\infty\}) = \{x : f(x) > a, \forall a \in \mathbb{R}\} = \{x : f(x) > a, \forall a \in Q\} = \bigcup_{a \in \mathbb{Q}} \{x : f(x) > a\}$$

is a countable union of measurable sets, therefore, measurable. ■

Proposition 2.1.1 $f = \chi_E$ is measurable if and only if $E \subseteq \mathbb{R}$ is a measurable set.

Proof. We note that $\{x : f(x) < \alpha\}$ will be one of $\mathbb{R}, E^c, \emptyset$, depending on whether $\alpha > 1, 0 < \alpha < 1, \alpha \leq 0$. Hence $f^{-1}([-\infty, \alpha))$ is a measurable set for all α if and only if E is measurable. ■

Definition 2.1.2 — Simple Function. A simple function is a function of the form $f = \sum_{i=1}^{N} \alpha_i \chi_{E_i}$ where $E_i \subseteq \mathbb{R}$ are measurable sets and $a_i \in \mathbb{R}$.

R A common example of a simple function can be a step function. But in general, the simple functions are not simple nor naive.

Exercise 2.2 Use the definition of a measurable function to check that every simple function is measurable.
Solution: Let $f = \sum_{i=1}^{N} a_i \chi_{E_i}$ be a simple function where $E_i \subseteq \mathbb{R}$ are measurable and $a_i \in \mathbb{R}$.

Claim: We can even write it as a linear combination of disjoint characteristic functions

Proof. For each N-tuple b, we can define

$$T_b(i) = \begin{cases} E_i & b_i = 0 \\ E_i^c & b_i = 1 \end{cases}$$

then,

$$T_b = \bigcap_{i=1}^{N} T_b(i)$$

is a measurable set by finite intersection of measurable sets. Also, T_b is pairwise disjoint for all possible b N-tuples. Then, we can rewrite

$$f = \sum_b \left(\sum_{i=1}^{N} b_i a_i \right) \chi_{T_b}$$

■

Then, for any $b \in \mathbb{R}$,

$$\{x : f(x) < b\} = \left\{ x : \sum_{i=1}^{N} a_i \chi_{E_i}(x) < b \right\} = \bigcup_{i=1}^{N} \{x : a_i \chi_{E_i}(x) < b\}$$

thus, measurable. ■

2.2 Properties of Measurable Functions

Proposition 2.2.1 If f, g are real-valued, measurable functions (with the same domain), then so are $f \pm g, fg$ and f/g if $g \neq 0$.

Proof. We note that

$$\{x : (f+g)(x) < \alpha\} = \{x : f(x) < \alpha - g(x)\}$$

Then, there is $r \in \mathbb{Q}$ such that $f(x) < r < \alpha - g(x)$. Then,

$$\{x : (f+g)(x) < \alpha\} = \bigcup_{r \in \mathbb{Q}} \{x : f(x) < r\} \cap \{x : g(x) < \alpha - r\}$$

Since f, g are measurable, the sets $\{x : f(x) < r\}, \{x : g(x) < \alpha - r\}$ are measurable for each choice of α, r. Since we are taking a countable union of measurable sets, it is still measurable. Thus, $f + g$ is measurable. ■

Exercise 2.3 Prove the case fg and f/g if $g \neq 0$.
Solution: The fg case can be resolve using he following trick

$$fg = \frac{1}{2}\left[(f+g)^2 - f^2 - g^2\right]$$

We shall show that $1/g$ when $g \neq 0$ is measurable. Let $a \in \mathbb{R}$, then

$$\{x : 1/g(x) < a\} = \left\{x : g(x) > \frac{1}{a}\right\}$$

which is measurable if $a \neq 0$. If $a = 0$, then the set becomes $\{x : g(x) < 0\}$ which is still measurable. ■

R We can extend this to extended real line.

Proposition 2.2.2 — Measurability behaves well under limits. Suppose $\{f_n\}_n$ are measurable functions. Then, so are $\sup_n f_n, \inf_n f_n$. If $f_n \to f$ pointwise as $n \to \infty$, then f is measurable.

Proof. Note that

$$\left\{x : \sup_n f_n(x) > \alpha\right\} = \bigcup_{n=1}^{\infty} \{x : f_n(x) > \alpha\}$$

and

$$\left\{x : \inf_n f_n(x) < \alpha\right\} = \bigcup_{n=1}^{\infty} \{x : f_n(x) < \alpha\}$$

Thus, $\sup_n f_n, \inf_n f_n$ are two measurable funcitons.
If $f_n \to f$, then $f = \liminf f_n = \sup_n(\inf_{k\geq n} f_k)$, thus f is measurable. Similarly, for $\limsup f_n$. ■

Theorem 9 The positive, real-valued function $f : \mathbb{R} \to [0,\infty)$ is measurable if and only if there are simple functions ϕ_n, $n \in \mathbb{N}$, with $\phi_n \leq \phi_{n+1}$ and $\phi_n \to f$ pointwise.

Proof. One direction is clear: simple functions are measurable, then their pointwise limits are measurable.
Suppose f is measurable and for $k = 0, 1, \cdots, n2^n - 1$, let

$$E_{n,k} = f^{-1}([k2^{-n}, (k+1)2^{-n})) = \left\{x : k2^{-n} \leq f(x) < (k+1)2^{-n}\right\}$$

then

$$E_n = \bigcup_{k=1}^{n2^n-1} E_{n,k} = f^{-1}[0,n)$$

is measurable. Let

$$\phi_n(x) = \begin{cases} k2^{-n} & x \in E_{n,k} \\ n & x \notin E_n \end{cases} = \sum_{k=0}^{n2^n-1} \frac{k}{2^n}\chi_{E_{n,k}} + n\chi_{E_n^c}$$

note that ϕ_n are simple and the choice of partitioning the range into subintervals of with 2^{-n} ensures that $\phi_n \leq \phi_{n+1}$.
It remains to check that $\phi_n \to f$ pointwise. For x such that $f(x) < N$, $x \in E_n$ for all $n \geq N$, say $x \in E_{n,k}$ for the suitable choice of k. Our construction ensures that

$$f(x) \in [k2^{-n}, (k+1)2^{-n})$$

and $\phi_n(x) = k2^{-n}$. Then,

$$|f(x) - \phi_n(x)| < 2^{-n}, \forall n \geq N$$

■

2.2.1 Almost Everywhere

Definition 2.2.1 — Almost Everywhere. We say that $f = g$ almost everywhere, and write $f = g$ a.e. if $m\{x : f(x) \neq g(x)\} = 0$.

■ **Example 2.2** $\chi_{\mathbb{Q}} = 0$ a.e. since $m(\mathbb{Q}) = 0$. ■

Proposition 2.2.3 1. If $f = 0$ a.e., then f is measurable

Proof. For $a \in \mathbb{R}$, if $a \leq 0$, then

$$\{x : f(x) < a\} \subseteq \{x : f(x) \neq 0\}$$

is measurable with 0. If $a > 0$, then

$$\{x : f(x) > a\} \subseteq \{x : f(x) \neq 0\}$$

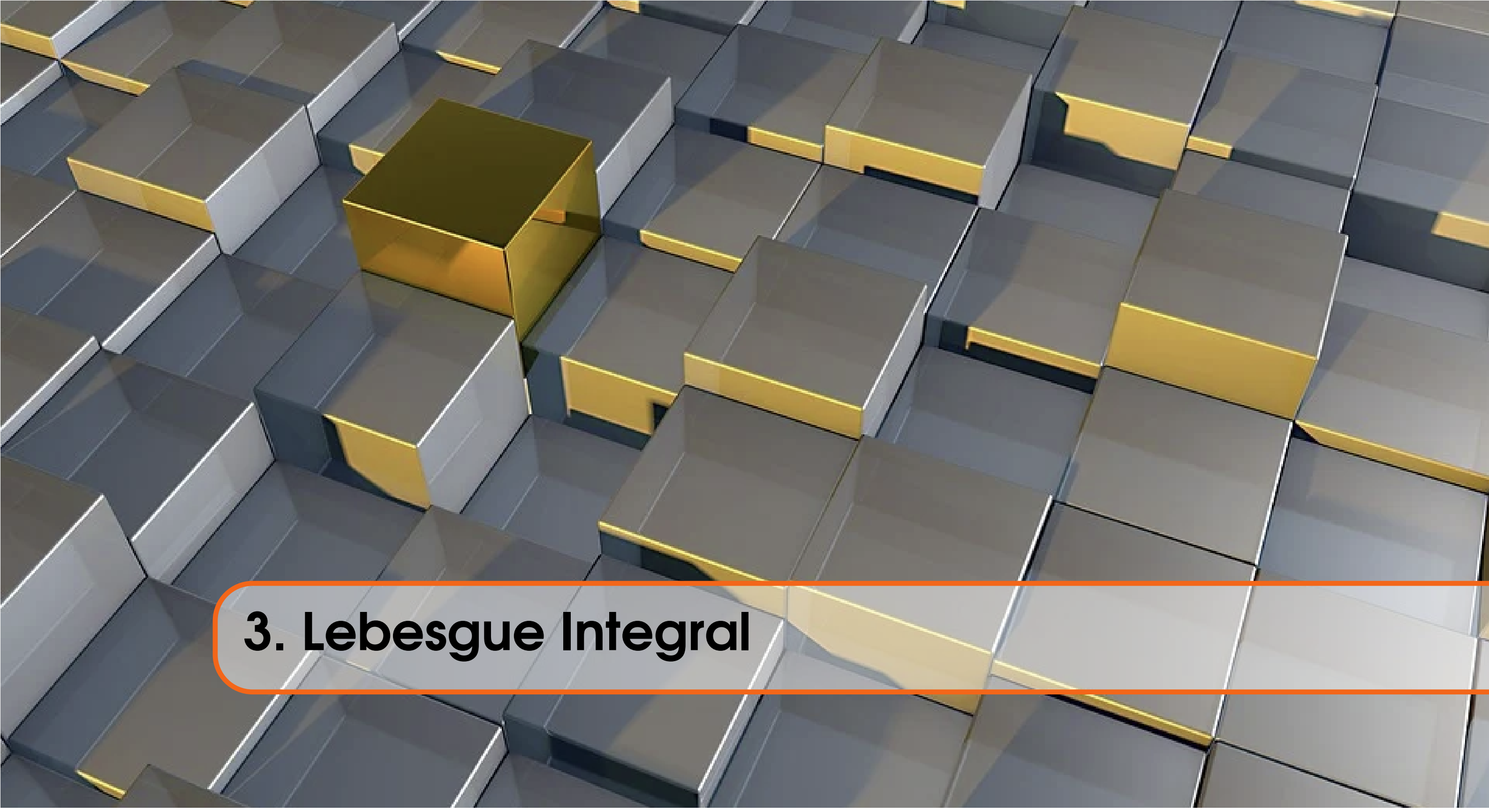
is measurable with 0. Thus, f is measurable. ■

2. If $f = g$ a.e., and f is measurable, then g is measurable

Proof. Let $h = f - g = 0$ a.e and f is measurable, by 1, we have

$$f - h = g$$

is measurable. ■



3. Lebesgue Integral

3.1 Definition of the Lebesgue Integral

Recall that a simple function $\phi = \sum_{k=1}^{N} a_k \chi_{I_k}$, where I_k are intervals and $a_j \in \mathbb{R}$,

$$R - \int_a^b \phi = \sum_k a_k l(I_k \cap [a,b])$$

> R It is certainly true that any simple function ϕ has a unique representation as $\sum_{k=1}^{N} a_k \chi_{E_K}$ where a_k are distinct and sets E_k are disjoint and their union is the whole real line.
> For convenience, we set convention that $0 \cdot \infty = 0$ instead of indeterminate.

Definition 3.1.1 — Lebesgue Integral of Simple Function. Assume $\phi \geq 0$ is a simple function with standard representation $\phi = \sum_{k=1}^{N} a_k \chi_{E_k}$. We define the Lebesgue integral of ϕ over the measurable set E as

$$\int_E \phi = \sum_{k=1}^{N} a_k m(E_k \cap E)$$

Exercise 3.1 Show that if ϕ also has a representation $\phi = \sum_{j=1}^{M} b_j \chi_{F_j}$ where the sets F_j are measurable, then $\int_E \phi = \sum_{j=1}^{M} b_j m(F_j \cap E)$
Solution: This is really just trying to rewrite F_j into disjoint union by finding distinct b_j values and them together using $\sigma-$additivity. ■

■ **Example 3.1**

1. If ϕ is a simple function and $E = [a,b]$, then

$$\int_E \phi = R - \int_a^b \phi$$

2. If $\phi = \chi_{\mathbb{Q}}$ and $E = [0,1]$, then as ϕ has standard representation $\phi = 1 \cdot \chi_{\mathbb{Q}} + 0 \cdot \chi_{\mathbb{Q}^c}$, we have

$$\int_E \phi = 1m(\mathbb{Q} \cap [0,1]) + 0m(\mathbb{Q}^c \cap [0,1]) = 0$$

■

Definition 3.1.2 — Lebesgue Integral of Non-negative Measurable Function. Suppose $f : E \subseteq \mathbb{R} \to [0,\infty]$ is measurable and E is a measurable set. Then, we define the Lebesgue integral of f over E as

$$\int_E f = \sup\left\{\int_E \phi : 0 \le \phi \le f, \phi \text{ simple}\right\}$$

Exercise 3.2 If $\phi = 0$ a.e., then $\int_E \phi = 0$ for any measurable set E.
Solution:
If $\phi = 0$ a.e., then $\{x : \phi(x) \neq 0\}$ has measure 0. Let $E \subseteq \mathbb{R}$ be any measurable set.

$$\int_E \phi = \sup\left\{\int_E \delta : 0 \le \delta \le \phi, \delta \text{ simple}\right\}$$

note that for any δ in the set above, each one of them are also equal to 0 a.e. since

$$\{x : \delta(x) \neq 0\} \subseteq \{x : \phi(x) \neq 0\}$$

then,

$$\int_E \delta = \sum_{k=1}^{N} a_k m(E_k \cap E) = 0$$

then, taking supremum, we have

$$\int_E \phi = 0$$

■

We can extend this idea to any type of function.

Definition 3.1.3 — Lebesgue Integral over $E \subseteq \mathbb{R}$. For measurable E and $f : E \to \mathbb{R}$, we define integral of f over E as

$$\int_E f = \int_E f^+ - \int_E f^-$$

provided this is not a $\infty - \infty$ form, where

$$f^+(x) = \begin{cases} f(x) & f(x) \ge 0 \\ 0 & \text{otherwise} \end{cases} = \max(f,0)$$

$$f^+(x) = \begin{cases} 0 & f(x) \ge 0 \\ -f(x) & \text{otherwise} \end{cases} = \max(-f,0)$$

R

1. It is clear that $f = f^+ - f^-$
2. f^+, f^- are measurable since we can decompose its domain into where the function is 0 and where it is positive.

Definition 3.1.4 — Lebesgue Integrable. We say the measurable function f is integrable on E if $\int_E |f| < \infty$. (Note that $|f| = f^+ + f^-$)

R Since $f^+, f^- \leq |f|$, by the monotonicity property of the integral of non-negative functions, $\int_E f^+$ and $\int_E f^-$ are finite for any integrable function f and hence $\int_E f$ is well defined.

Definition 3.1.5 — Lebesgue Integrable on $\mathbb{C}$. If $f : E \subseteq \mathbb{R} \to \mathbb{C}$, we say f is integrable if the both $\operatorname{Re} f, \operatorname{Im} f$ are integrable functions and then we define the Lebesgue integral of f over E as

$$\int_E f = \int_E \operatorname{Re} f + i \int_E \operatorname{Im} f$$

3.2 Properties of Lebesgue Integral

Proposition 3.2.1 — Properties of Lebesgue Integral.

1. **Monotonicity:** if $0 \leq f \leq g$, then $\int_E f \leq \int_E g$. In particular, if $|f|$ is bounded by M, then

$$\int_E |f| \leq \int_E M = Mm(E)$$

2. $\int_E f = \int_{\mathbb{R}} f\chi_E$. As a consequence, there is no loss in assuming $E = \mathbb{R}$ and in this case, we will write $\int f$.

Proof. First note that $f\chi_E$ is measurable.

(a) Simple case: suppose $f = \sum_{k=1}^N a_k \chi_{E_k}$, then

$$\begin{aligned}\int_E f &= \sum_{k=1}^N a_k m(E_k \cap E) = \int_{\mathbb{R}} \sum_{k=1}^N a_k \chi_{E_k \cap E} \\ &= \int_{\mathbb{R}} \left(\sum_{k=1}^N a_k \chi_{E_k} \right) = \int_{\mathbb{R}} f\chi_E\end{aligned}$$

(b) Positive case: let ϕ be simple function such that $0 \leq \phi \leq f$, then $\phi\chi_E$ is also a simple function with $0 \leq \phi\chi_E \leq f\chi_E$. Thus,

$$\int_E \phi = \int \phi\chi_E \leq \sup\left\{ \int_E \psi : 0 \leq \psi \leq f\chi_E \right\} = \int f\chi_E$$

As this is true for all such ϕ, we have

$$\int_E f = \sup\left\{ \int_E \phi : 0 \leq \phi \leq f \right\} \leq \int f\chi_E$$

On the other hand, if ψ is a simple function with $0 \leq \psi \leq f\chi_E$, then also $0 \leq \psi \leq f\chi_E \leq f$ and note that $\psi\chi_E = \psi$. Thus,

$$\int \psi = \int \psi\chi_E = \int_E \psi \leq \sup\left\{ \int_E \phi : 0 \leq \phi \leq f \right\} = \int_E f$$

Therefore,

$$\int f\chi_E = \sup\left\{ \int \psi : 0 \leq \psi \leq f\chi_E \right\} \leq \int_E f$$

Thus, $\int f\chi_E = \int_E f$.

(c) If f is a real-valued and integrable function over E, then since $(f\chi_E)^{\pm} = f^{\pm}\chi_E$, the previous step implies $\int (f\chi_E)^{\pm} = \int f^{\pm}\chi_E = \int_E f^{\pm}$. It follows that $f\chi_E$ is integrable over $\mathbb{R}$ and

$$\int f\chi_E = \int f^+\chi_E - \int f^-\chi_E = \int_E f^+ - \int_E f^- = \int_E f$$

(d) Complex case: really similar to the real-valued case.

∎

3. If $m(E) = 0$, then $\int_E f = 0$ (even if $f = \infty$ on E)

Proof. If $f = \sum a_k \chi_{E_k}$ is simple, then
$$\int_E f = \int_E \sum a_k \chi_{E_k} = \sum a_k m(E_k \cap E) = 0$$
If $f \geq 0$, then taking supremum of a bunch of zero is still 0. Real-valued and complex-valued cases are similar. ∎

4. For all scalar $\alpha \in \mathbb{C}$, $\int \alpha f = \alpha \int f$

Proof. Suppose h is real-valued and integrable.

(a) For $\alpha \geq 0$:

i. For simple function's case, say $h = \sum_{i=1}^N a_i \chi_{E_i}$, for any $E \supseteq \mathbb{R}$,
$$\begin{aligned}\int_E \alpha h &= \int_E \alpha \sum_{i=1}^N a_i \chi_{E_i} \\ &= \sum_{i=1}^N \alpha a_i m\{E_i \cap E\} \\ &= \alpha \sum_{i=1}^N a_i m\{E_i \cap E\} \\ &= \alpha \int_E h\end{aligned}$$

ii. For positive function's case, note if ψ is a simple function with $0 \leq \psi \leq h$, note that $\alpha\psi$ is still a simple function such that $0 \leq \alpha\psi \leq \alpha h$. Then,
$$\int \alpha h = \sup\left\{\int \phi : 0 \leq \phi \leq \alpha h\right\} \geq \int \alpha\psi = \alpha \int \psi$$
$$\alpha \int h = \alpha \sup\left\{\int \psi : 0 \leq \psi \leq h\right\} = \sup\left\{\alpha \int \psi : 0 \leq \psi \leq h\right\} \leq \int \alpha h$$
on the other hand, since $\alpha \geq 0$, we have
$$\begin{aligned}\alpha \int h &= \alpha \sup\left\{\int \psi : 0 \leq \psi \leq h\right\} = \sup\left\{\alpha \int \psi : 0 \leq \psi \leq h\right\} \\ &= \sup\left\{\int \alpha\psi : 0 \leq \alpha\psi \leq \alpha h\right\} \leq \int \alpha h\end{aligned}$$
Thus, $\int \alpha h = \alpha \int h$.

iii. for h being real-valued, it is clear that given $\alpha > 0$,
$$\begin{aligned}\int \alpha h &= \int (\alpha h)^+ - \int (\alpha h)^- \\ &= \int \max(\alpha h, 0) - \int \max(-\alpha h, 0) \\ &= \int \alpha \max(h, 0) - \int \alpha \max(-h, 0) \\ &= \alpha \int h^+ - \alpha \int h^- \\ &= \alpha\left(\int h^+ - \int h^-\right) = \alpha \int h\end{aligned}$$

(b) For $\alpha < 0$, we shall show $\alpha = -1$ case, which will be enough.

$$-\int h = -\left(\int h^+ - \int h^-\right) = \int h^- - \int h^+ = \int \max(-h,0) - \int \max(h,0)$$

$$= \int (-h)^+ - \int (-h)^- = \int -h$$

■

5. **Triangle Inequality:**

$$\left|\int f\right| \leq \int |f|$$

Proof. Choose α with $|\alpha| = 1$ satisfying $|\int f| = \alpha \int f$, then

$$\left|\int f\right| = \int \alpha f = \int \mathrm{Re}\, \alpha f + i \int \mathrm{Im}\, \alpha f = \int \mathrm{Re}\, \alpha f$$

But $\mathrm{Re}\, \alpha f \leq |\alpha f| = |f|$, then by monotonicity,

$$\left|\int f\right| = \int \mathrm{Re}\, \alpha f \leq \int |\alpha f| = \int |f|$$

■

6. For all ϕ, ψ simple, we have

$$\int \phi + \psi = \int \phi + \int \psi$$

Proof. (a) Let ϕ, ψ be simple and positive, say

$$\phi = \sum_{i=1}^{N} a_i \chi_{E_i}$$

$$\psi = \sum_{k=1}^{M} b_k \chi_{F_k}$$

then,

$$\phi + \psi = \sum_{i=1}^{N} a_i \chi_{E_i} + \sum_{k=1}^{M} b_k \chi_{F_k}$$

Let $\{c_r\}$ be the unique set of the finite set of values $\{a_i\}_{i=1}^N$ and $\{b_k\}_{k=1}^M$. For each c_r, we can take the index i such that $a_i = c_r$ and the index k such that $b_k = c_r$ and consider $E_i \cup F_k$ (if we cannot find any one of E_i, F_k, let pick $\emptyset$). Let $V_{r,1} = E_i \cap F_k$ and pair it up with the value $2c_r = c_{r,1}$, while $V_{r,2} = (E_i \cup F_k) \backslash V_{r,1}$ is paired up with $c_r = c_{r,2}$. Note that $V_{r,1}, V_{r,2}, \forall r$ are measurable and disjoint. Since there are only finitely many $\{c_i\}$, we can perform this procedure for all of them. Then, we have the following simple function

$$\phi + \psi = \sum_{r=1}^{L} c_{r,1} V_{r,1} + c_{r,2} V_{r,2}$$

for any $E \subseteq \mathbb{R}$

$$\begin{aligned}
\int_E \phi + \psi &= \int_E \sum_{r=1}^{L} c_{r,1}V_{r,1} + c_{r,2}V_{r,2} \\
&= \sum_{r=1}^{L} c_{r,1}m(V_{r,1} \cap E) + c_{r,2}m(V_{r,2} \cap E) \\
&= \sum_{r=1}^{L} 2c_r m((E_i \cap F_k) \cap E) + c_r m(((E_i \cup F_k) \setminus (E_i \cap F_k)) \cap E) \\
&= \sum_{i=1}^{N} a_i m(E_i \cap E) + \sum_{k=1}^{M} b_k m(F_k \cap E)
\end{aligned}$$

Assignment 2 result and construction

$$\begin{aligned}
&= \int \sum_{i=1}^{N} a_i \chi_{E_i} + \int \sum_{k=1}^{M} b_k \chi_{F_k} \\
&= \int_E \phi + \int_E \psi
\end{aligned}$$

(b)

Lemma 3.3 If $f_1, f_2 \geq 0$ and integrable, and $f = f_1 - f_2$, then f is integrable and $\int f = \int f_1 - \int f_2$.

Proof. Consider

$$\int |f| = \int |f_1 - f_2| \leq \int |f_1| + |f_2|$$

since f_1, f_2 are integrable, f_1, f_2 are measurable and $|f_1|, |f_2| \geq 0$ and measurable by continuous composition. Thus,

$$\int |f_1| + |f_2| = \int |f_1| + \int |f_2|$$

then,

$$\int |f| \leq \int |f_1| + \int |f_2| < \infty$$

Therefore, f is integrable.
Then,

$$\int f = \int f^+ - \int f^-$$

where

$$f^+ = \max(f, 0) = \max(f_1 - f_2, 0) = \max(f_2 - f_1, 0) - (f_2 - f_1) = f^- - f_2 + f_1$$

then,

$$f^+ + f_2 = f^- + f_1 \implies \int f^+ + f_2 = \int f^- + f_1 \implies \int f^+ + \int f_2 = \int f^- + \int f_1$$

since these are all non-negative measurable functions. Then,

$$\int f = \int f^+ - \int f^- = \int f_1 - \int f_2$$

■

Let f, g be complex-valued and integrable, say

$$f = f_R + i f_I, f_R = Re(f), f_I = Im(f)$$

$$g = g_R + i g_I, g_R = Re(g), g_I = Im(g)$$

then, by definition, f_R, f_I, g_R, g_I are all integrable real-valued functions. Then,

$$\begin{aligned}
\int f+g &= \int Re(f+g) + i\int Im(f+g) \\
&= \int Re(f) + Re(g) + i\int Im(f) + Im(g) \\
&= \int f_R + g_R + i\int f_I + g_I \\
&= \int f_R^+ - f_R^- + g_R^+ - g_R^- + i\int f_I^+ - f_I^- + g_I^+ - g_I^- \\
&= \int (f_R^+ + g_R^+) - (f_R^- + g_R^-) + i\int (f_I^+ + g_I^+) - (f_I^- + g_I^-) \\
&= \int (f_R^+ + g_R^+) - \int (f_R^- + g_R^-) + i\left[\int (f_I^+ + g_I^+) - \int (f_I^- + g_I^-)\right] && \text{by lemma} \\
&= \int f_R^+ + \int g_R^+ - \int f_R^- - \int g_R^- + i\left[\int f_I^+ + \int g_I^+ - \int f_I^- - \int g_I^-\right] \\
&\text{by proved proposition} \\
&= \int f_R^+ - \int f_R^- + i\left[\int f_I^+ - \int f_I^-\right] + \int g_R^+ - \int g_R^- + i\left[\int g_I^+ - \int g_I^-\right] \\
&= \int f_R + i\int f_I + \int g_R + i\int g_I && \text{by lemma} \\
&= \int f + \int g
\end{aligned}$$

■

7. **Translation Invariance of Lebesgue Integral:** for fixed $y \in \mathbb{R}$,

$$\int_{\mathbb{R}} f(x+y)dx = \int_{\mathbb{R}} f(x)dx$$

Proof. We follow the sketch given in the lecture notes.

Claim: For fixed $y \in \mathbb{R}$, if f is measurable, then $f(x+y)$ is also measurable
Proof. For $\alpha \in \mathbb{R}$,

$$\{x : f(x) < \alpha\}$$

is measurable, then by Assignment 1, we know that

$$\{x : f(x) < \alpha\} + y = \{x : f(x+y) < \alpha\}$$

is also measurable. Thus, $x \mapsto f(x+y)$ is measurable. ■

We observe that the translation invariance of Lebesgue measurable proved in Assignment 1 and the Lebesgue integral of a characteristic function give us: for any measurable set

$$\int \chi_E(x+y)dx = \int \chi_{E-y}(x)dx = m(E-y) = m(E) = \int \chi_E(x)dx$$

this is true since

$$\chi_E(x+y) = \begin{cases} 1 & x+y \in E \\ 0 & x+y \notin E \end{cases} = \begin{cases} 1 & x \in E-y \\ 0 & x \notin E-y \end{cases} = \chi_{E-y}(x)$$

(a) If f is simple with standard representation

$$f=\sum_{i=1}^{N} a_i \chi_{E_i}$$

then, by property 4 and 6, we have

$$\int f(x+y)dx = \int \left[\sum_{i=1}^{N} a_i \chi_{E_i}(x+y)\right] dx = \sum_{i=1}^{N} a_i \int \chi_{E_i}(x+y)dx$$

$$= \sum_{i=1}^{N} a_i \int \chi_{E_i}(x)dx = \int \left[\sum_{i=1}^{N} a_i \chi_{E_i}(x)\right] dx = \int f(x)dx$$

(b) If f is positive, consider

$$\begin{aligned}\int f(x+y)dx &= \sup\left\{\int \phi(x+y)dx : 0 \le \phi(x+y) \le f(x+y)\right\}\\ &= \sup\left\{\int \phi(x)dx : 0 \le \phi(x+y) \le f(x+y)\right\} && \text{by simple function}\\ &= \sup\left\{\int \phi(x)dx : 0 \le \phi(x) \le f(x)\right\}\\ & 0 \le \phi(x) \le f(x) \iff 0 \le \phi(x+y) \le f(x+y), \forall x \in \mathbb{R}\\ &= \int f(x)dx\end{aligned}$$

(c) suppose f is integrable and real-valued, so f^+, f^- are all integrable and positive.

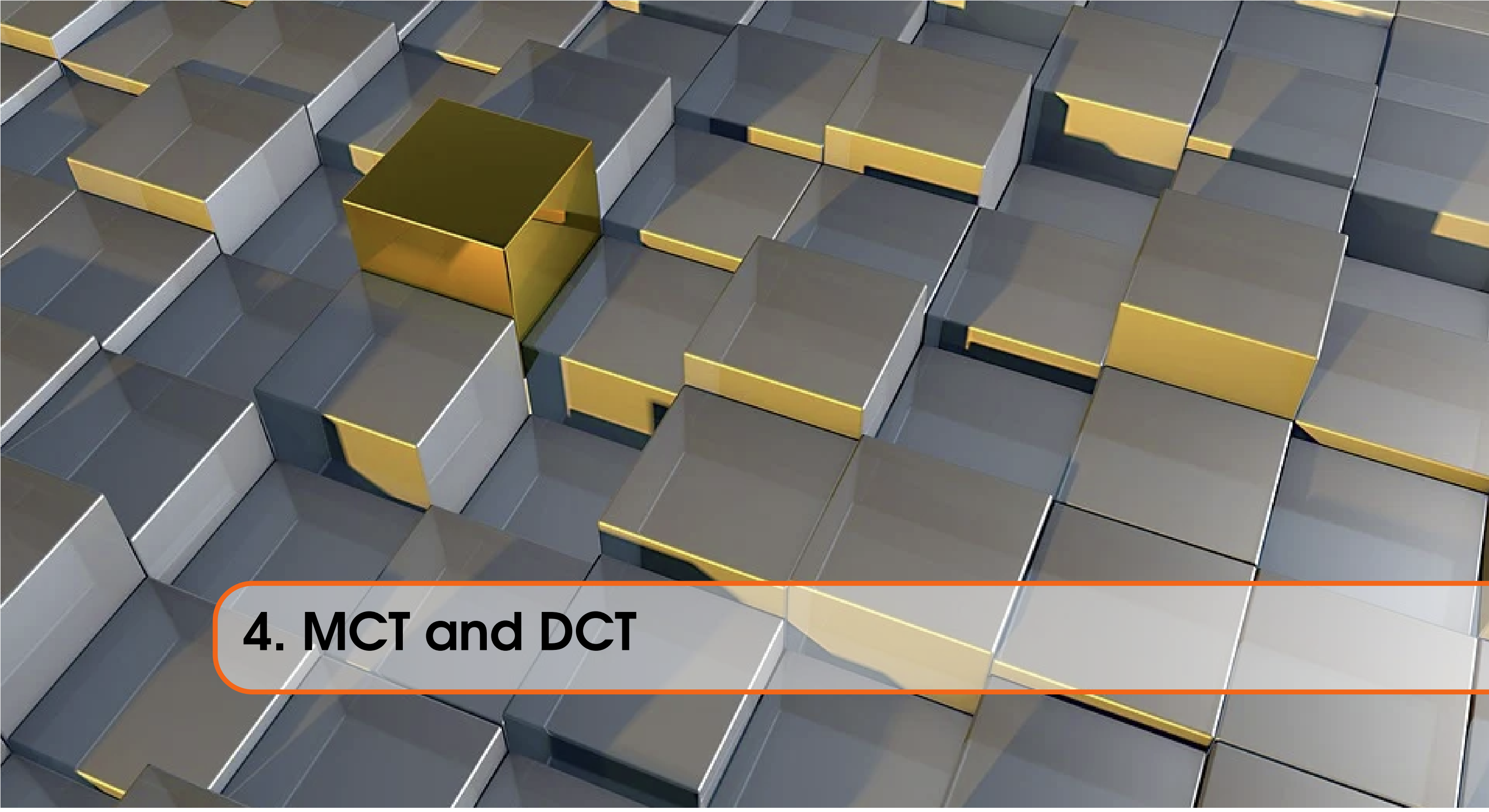
$$\begin{aligned}\int f(x+y)dx &= \int f^+(x+y)dx - \int f^-(x+y)dx\\ &= \int f^+(x)dx - \int f^-(x)dx\\ &= \int f(x)dx\end{aligned}$$

(d) The case for f is integrable and complex valued, we have

$$\begin{aligned}\int f(x+y)dx &= \int Re(f(x+y))dx + i\int Im(f(x+y))dx\\ &= \int \mathrm{Re}(f(x))dx + i\int \mathrm{Im}(f(x))dx\end{aligned}$$

since $Re(f), Im(f)$ are integrable real-valued functions.

■

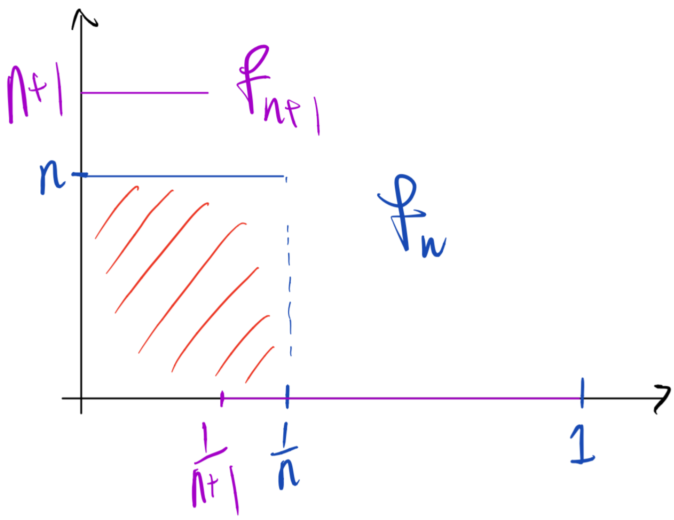


4. MCT and DCT

4.1 Monotone Convergence Theorem and Fatou's Lemma

We have seen that Riemann integral does not work quite while with limit, unfortunately, it will be generally the case for Lebesgue integral again.

■ **Example 4.1** Define f_n on $[0,1]$ by $f_n(x) = n$ for $x \in \left(0, \frac{1}{n}\right)$ and $f_n(x) = 0$ else. See the picture.



Then, f_n is a simple (step) function and $f_n \to 0$ positively. However,

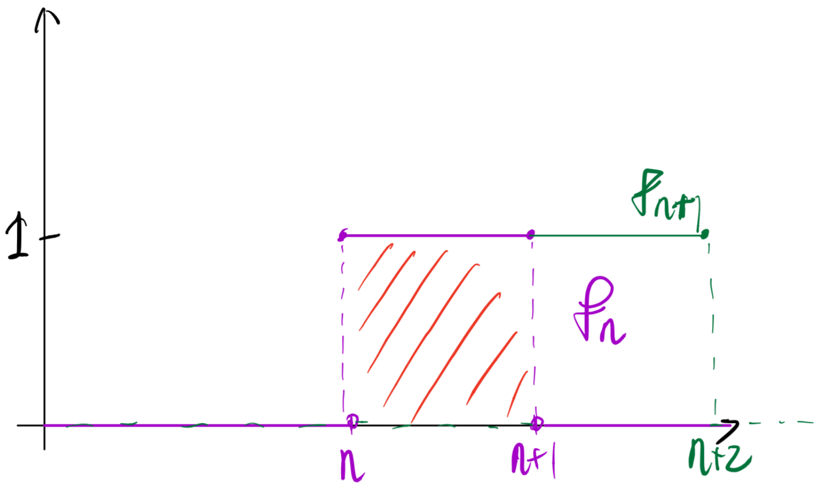
$$\int_{[0,1]} f_n = 1, \forall n \in \mathbb{N}$$

while

$$\int_{[0,1]} \lim_n f_n = \int 0 = 0$$

■

■ **Example 4.2** Define f_n on $\mathbb{R}$ by $f_n(x) = \chi_{[n,n+1]}$. See the picture.



Again, $f_n \to 0$ positive and $\int f_n = 1$ and $\int \lim f_n = 0$.

■

Theorem 10 — (Lebesgue) Monotone Convergence Theorem. Suppose $f_n \geq 0$ and measurable. If $f_n(x) \leq f_{n+1}(x)$ for all x and n, and $f_n \to f$ pointwise, then

$$\lim_{n\to\infty} \int_E f_n = \int_E \lim_n f_n = \int_E f$$

for any measurable set E.

Proof. Since each f_n is measurable and converging to f pointwise, by a previous theorem, we know that f is also measurable. As $f_n \uparrow f$, we have $f_n \leq f(x)$ for all n and x, and thus,

$$\int_E f_n \leq \int_E f$$

Since the sequence of integrals $(\int_E f_n)_n$ is also increasing, it must have a limit (possibly $+\infty$). Thus,

$$\lim_n \int_E f_n \leq \int_E f$$

we want to show equality.
Let $\phi \leq f$ be a simple function and $\alpha \in (0,1)$. Let

$$A_n = \{x \in E : f_n(x) \geq \alpha\phi(x)\} = (f - \alpha\phi)^{-1}([0,\infty]) \cap E$$

these are measurable sets and since $f_n \leq f_{n+1}$ we have $A_n \subseteq A_{n+1}$ for each n (nested!).
Suppose $x \in E$. If $\phi(x) = 0$, then we clearly have $f_n(x) \geq \alpha\phi(x)$ for all n and hence $x \in A_n$ for all n.
If $\phi(x) \neq 0$, then $\alpha\phi(x) < \phi(x) \leq f(x)$. Since $f_n \to f$, we have $f_n \geq \alpha\phi$ eventually. In any case $x \in \cup A_n$. Since $A_n \subseteq E$, we have $E = \cup A_n$. Then, by monotonicity property of the integral

implies that

$$\alpha \int_{A_n} \phi = \int_{A_n} \alpha\phi \leq \int_{A_n} f_n = \int f_n \chi_{A_n} \leq \int f_n \chi_E = \int_E f_n \to \lim_n \int_E f_n$$

Applying Lemma 4.2, we have

$$\alpha \int_E \phi = \alpha \lim_n \int_{A_n} \phi \leq \lim_n \int_E f_n$$

but $\int_E f = \sup_{\text{simple } \phi \leq f} \int_E \phi$, thus,

$$\alpha \int_E \leq \lim_n f_n, \forall \alpha < 1$$

Letting $\alpha \to 1$, this completes the proof. ■

R

1. We allow f_n and f to take $+\infty$ and the possibility of $\lim_n \int_E f_n = +\infty$.
2. The extra hypothesis here is that the sequence f_n converges **up** to f, therefore "increasing".

R This version of the MCT does not necessarily hold for the Riemann integral.
A counter example is the sequence of functions

$$f_n = \begin{cases} 1 & x \in \{r_1, r_2, \cdots, r_n\} \\ 0 & \text{otherwise} \end{cases}$$

where $\mathbb{Q} = \{r_n\}_n$, then clearly $f_n \uparrow \chi_{\mathbb{Q}}$. The functions are all Riemann integrable and $R - \int_0^1 f_n = 0$ but the limit function $\chi_{\mathbb{Q}}$ is not Riemann integrable, so is the integrable of the limit function. But it will hold in the notion of Lebesgue integral.

We shall prove the special case of simple function first, it will referred in the actual proof above. The main idea of the proof is the continuity of measure.

Lemma 4.2 Let $\phi \geq 0$ be a simple function assume A_n, $n = 1, 2, \cdots$ are measurable sets with $A_n \subseteq A_{n+1}$ for each n. Put $A = \bigcup_{n=1}^{\infty} A_n$, then

$$\lim_n \int_{A_n} \phi = \lim_n \int \phi \chi_{A_n} = \int_A \phi = \int \phi \chi_A$$

(note that $\phi \chi_{A_n} \uparrow \phi \chi_A$ by nested sets)

Proof. Assume $\phi = \sum_{i=1}^{N} a_i \chi_{E_i}$, where the sets E_i are measurable. Then,

$$\int_{A_n} \phi = \sum_{i=1}^{N} a_i (E_i \cap A_n)$$

note that for each i, $E_i \cap A_n \subseteq E_i \cap A_{n+1}$ and also $\cup_n (E_i \cap A_n) = E_i \cap A$. By the continuity of measure, $m(E_i \cap A_n) \to m(E_i \cap A)$ as $n \to \infty$. Hence,

$$\int_{A_n} \phi \to \sum_{i=1}^{N} a_i m(E_i \cap A) = \int_A \phi$$

■

R The MCT will be helpful to prove the general case of the linearity of the integral.

Proposition 4.2.1 If $f, g \geq 0$ are measurable, then

$$\int_E f + g = \int_E f + \int_E g$$

Proof. Given the result for simple function. For $f, g \geq 0$, by a theorem from previous part, we have $\phi_n \uparrow f$ and $\psi_n \uparrow g$. Hence, also $\phi_n + \psi_n \uparrow f + g$. By the MCT, we have

$$\int_E f + g = \lim_n \int_E \phi_n + \psi_n = \lim_n \left(\int_E \phi_n + \int_E \psi_n \right) = \int_E f + \int_E g$$

■

Exercise 4.1 If $f \geq 0$ and measurable, then for any measurable set E, we have

$$\int f = \int_E f + \int_{E^c} f$$

Solution: Note that

$$\int_E f + \int_{E^c} f = \int f \chi_E + \int f \chi_{E^c}$$

$f\chi_E, f\chi_{E^c} \geq 0$ are both measurable functions. By the proposition above, we know that

$$\int f\chi_E + \int f\chi_{E^c} = \int f\chi_E + f\chi_{E^c} = \int_E f$$

■

Proposition 4.2.2 Show that if f, g are integrable, then $\int f + g = \int f + \int g$.

Proof.

Lemma 4.3 If $f_1, f_2 \geq 0$ and integrable, and $f = f_1 - f_2$, then f is integrable and $\int f = \int f_1 - \int f_2$.

Proof. Consider

$$\int |f| = \int |f_1 - f_2| \leq \int |f_1| + |f_2|$$

since f_1, f_2 are integrable, f_1, f_2 are measurable and $|f_1|, |f_2| \geq 0$ and measurable by continuous composition. Thus,

$$\int |f_1| + |f_2| = \int |f_1| + \int |f_2|$$

then,

$$\int |f| \leq \int |f_1| + \int |f_2| < \infty$$

Therefore, f is integrable.
Then,

$$\int f = \int f^+ - \int f^-$$

where

$$f^+ = \max(f, 0) = \max(f_1 - f_2, 0) = \max(f_2 - f_1, 0) - (f_2 - f_1) = f^- - f_2 + f_1$$

then,

$$f^+ + f_2 = f^- + f_1 \Longrightarrow \int f^+ + f_2 = \int f^- + f_1 \Longrightarrow \int f^+ + \int f_2 = \int f^- + \int f_1$$

since these are all non-negative measurable functions. Then,

$$\int f = \int f^+ - \int f^- = \int f_1 - \int f_2$$

■

Let f, g be complex-valued and integrable, say

$$f = f_R + if_I, f_R = Re(f), f_I = Im(f)$$
$$g = g_R + ig_I, g_R = Re(g), g_I = Im(g)$$

then, by definition, f_R, f_I, g_R, g_I are all integrable real-valued functions. Then,

$$\begin{aligned}
\int f+g &= \int Re(f+g) + i\int Im(f+g) \\
&= \int Re(f)+Re(g) + i\int Im(f)+Im(g) \\
&= \int f_R + g_R + i\int f_I + g_I \\
&= \int f_R^+ - f_R^- + g_R^+ - g_R^- + i\int f_I^+ - f_I^- + g_I^+ - g_I^- \\
&= \int (f_R^+ + g_R^+) - (f_R^- + g_R^-) + i\int (f_I^+ + g_I^+) - (f_I^- + g_I^-) \\
&= \int (f_R^+ + g_R^+) - \int (f_R^- + g_R^-) + i\left[\int (f_I^+ + g_I^+) - \int (f_I^- + g_I^-)\right] && \text{by lemma} \\
&= \int f_R^+ + \int g_R^+ - \int f_R^- - \int g_R^- + i\left[\int f_I^+ + \int g_I^+ - \int f_I^- - \int g_I^-\right] && \text{by proved proposition} \\
&= \int f_R^+ - \int f_R^- + i\left[\int f_I^+ - \int f_I^-\right] + \int g_R^+ - \int g_R^- + i\left[\int g_I^+ - \int g_I^-\right] \\
&= \int f_R + i\int f_I + \int g_R + i\int g_I && \text{by lemma} \\
&= \int f + \int g
\end{aligned}$$

■

Corollary 4.3.1 Let $f \geq 0$ and measurable, then

$$\int_{-n}^{n} f \to \int_{\mathbb{R}} f$$

Proof. Let $f_n = f\chi_{[-n,n]}$ be a sequence of measurable and positive functions, and also $f_n \uparrow f$ for all x. By the MCT,

$$\int_{-n}^{n} f = \int_{\mathbb{R}} f_n \to \int_{\mathbb{R}} f$$

■

Exercise 4.2 Prove the above corollary if f is just integrable.
Solution: ■

Lemma 4.4 — Fatou's Lemma. For $f_n \geq 0$ and measurable, we have

$$\int \liminf_n f_n \leq \liminf_n \int f_n$$

Proof. Let $g_n = \inf_{k \geq n} f_k$. Then g_n is measurable and $g_n \leq f_k$ for all $n \leq k$. Thus, $\int g_n \leq \int f_k$ for all $n \leq k$ and that implies

$$\int g_n \leq \inf_{k \geq n} \int f_k$$

Moreover, (g_n) is increasing, the $\lim_n \int g_n$ exists (possibly $+\infty$). Likewise, the sequence $(\inf_{k \geq n} \int f_k)_n$ is increasing in n and hence has a limit. Thus,

$$\lim_n \int g_n \leq \lim_n \left(\inf_{k \geq n} \int f_k \right) = \liminf_n \int f_n$$

Let $F(x) = \liminf_n f_n(x)$. Note that by definition, $g_n \uparrow F(x)$, by MCT,

$$\lim_n \int g_n = \int F = \int \liminf_n f_n$$

We are done by this point.
SANDWICH Memorization: integral signs as bread and liminfs as hams. ■

■ **Example 4.3** Suppose $f_n \to f$ pointwise and $\int |f_n| \leq 1$ for all n. By Fatou's lemma,

$$\int |f| \underset{|f_n| \to |f|}{=} \int \liminf |f_n| \underset{\text{Fatou}}{\leq} \liminf \int |f_n| \leq 1$$

Then, f is also integrable. ■

4.5 Dominated Convergence Theorem

Theorem 11 — Lebesgue Dominated Convergence Theorem. Suppose f_n are measurable functions with $f_n \to f$ pointwise. **Assume there is an integrable function** g **such that** $|f_n(x)| \leq g(x)$ **for all** x, n. Then,

$$\int \lim_n f_n = \int f = \lim_n \int f_n$$

Moreover (stronger result),

$$\int |f_n - f| \to 0$$

Proof. We first note that since $|f_n(x)| \leq g(x)$, we have

$$\int |f_n| \leq \int g < \infty$$

as g is integrable. Also, by previous example, we have f is also integrable and measurable. Consider $2g - |f - f_n|$. Since $|f_n| \leq g$, we have $|f| \leq g$. Thus, $2g - |f - f_n| \geq 0$. Since g is integrable and therefore measurable, we have $2g - |f - f_n|$ being measurable. Then,

$$\begin{aligned}
\int 2g &\underset{|f - f_n| \to 0}{=} \int \liminf_n (2g - |f - f_n|) \\
&\leq \liminf_n \int 2g - |f - f_n| \\
&= \int 2g + \liminf_n - \int |f - f_n| \\
&= \int 2g - \limsup_n \int |f - f_n|
\end{aligned}$$

Since g is integrable, we have

$$\limsup_n \int |f - f_n| \leq 0$$

But

$$0 \leq \liminf_n \int |f - f_n| \leq \limsup_n \int |f - f_n| \leq 0$$

In particular, the equality holds throughout and

$$\lim_n \int |f_n - f| = 0$$

We note that

$$\left| \int f_n - \int f \right| = \left| \int f_n - f \right| \leq \int |f_n - f| \to 0$$

Hence,

$$\int f_n \to \int f$$

■

4.6 Relationship Between Riemann and Lebesgue Integrals

We have seen Riemann integrals and Lebesgue integrals agree on simple functions. But this is actually true in general.

Theorem 12 If f is Riemann integrable over $[a,b]$, then f is integrable (Lebesgue) and the values coincide.

Proof. Let f be Riemann integrable, then $|f(x)| \leq C$ for all $x \in [a,b]$. Let E be the set of discontinuity of f over $[a,b]$ which is of measure 0. Note that

$$f = f\chi_E + f\chi_{E^c}$$

and $f\chi_E = 0$ a.e. and $f\chi_{E^c}$ is measurable since it is continuous. Thus, f is measurable by a proposition. Since

$$\int_{[a,b]} |f| \leq \int_{[a,b]} C \leq C(b-a)$$

thus, f is Lebesgue integrable.
Now, take any partition $P = a = x_0 < x_1 < \cdots < x_n = b$, and let

$$M_i = \sup\{f(x) : x \in [x_{i-1}, x_i]\}$$

$$m_i = \inf\{f(x) : x \in [x_{i-1}, x_i]\}$$

Then, using simple function integral defintion,

$$U(f,P) = \sum_{i=1}^{n} M_i(x_i - x_{i-1} = \int_{[a,b]} \sum_{i=1}^{n} M_i \chi_{[x_{i-1},x_i)}$$

$$L(f,P) = \sum_{i=1}^{n} m_i(x_i - x_{i-1} = \int_{[a,b]} \sum_{i=1}^{n} m_i \chi_{[x_{i-1},x_i)}$$

We have

$$\sum_{i=1}^{n} M_i \chi_{[x_{i-1},x_i)} \geq f \geq \sum_{i=1}^{n} m_i \chi_{[x_{i-1},x_i)}$$

the by monotonicity,

$$U(f,P) = \int_{[a,b]} \sum_{i=1}^{n} M_i \chi_{[x_{i-1},x_i)} \geq \int_{[a,b]} f \geq \int_{[a,b]} \sum_{i=1}^{n} m_i \chi_{[x_{i-1},x_i)} = L(f,P)$$

Since f is Riemann integrable,

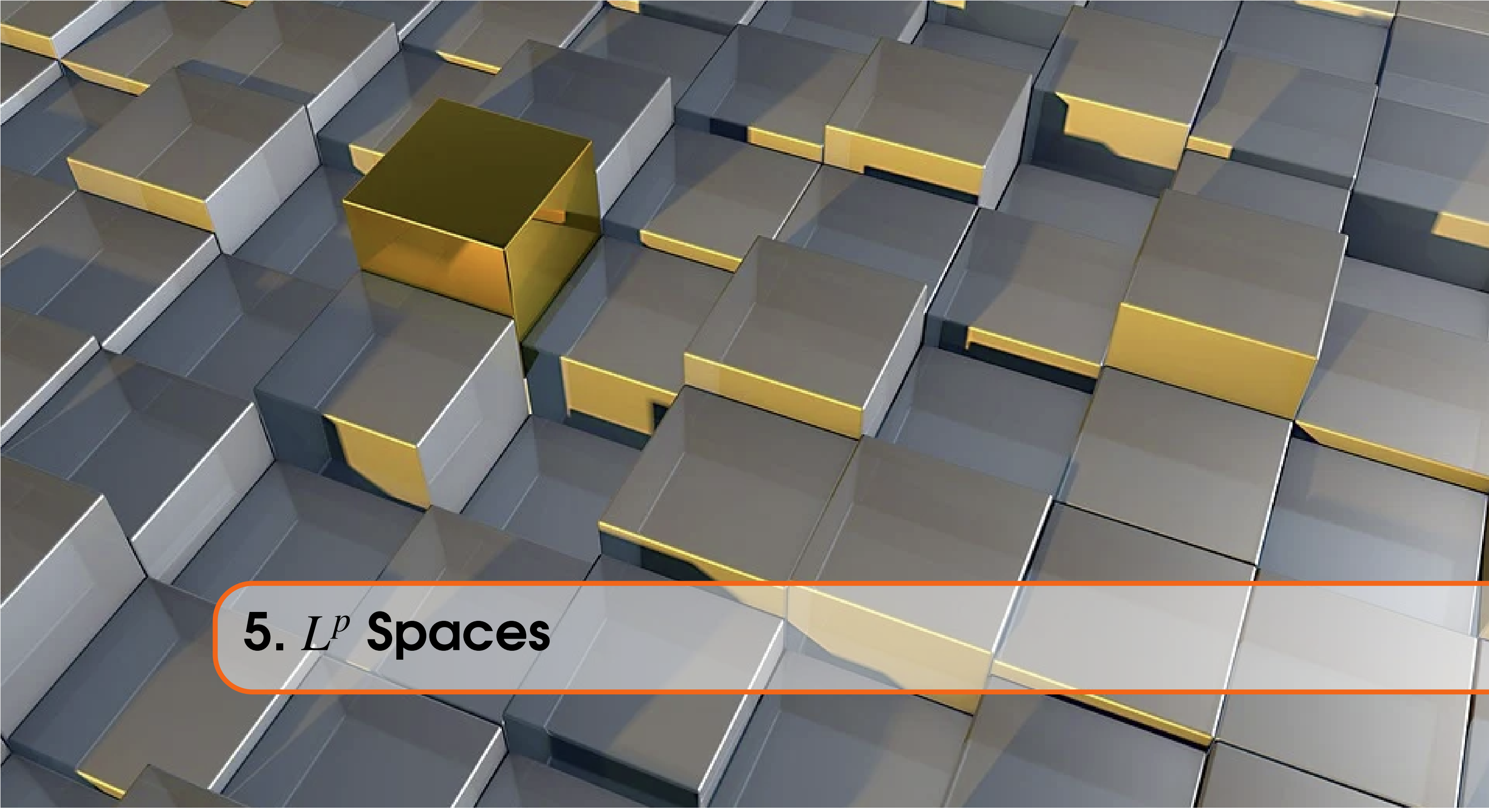
$$\inf_P U(f,P) = \sup_P L(f,P) = R - \int_a^b f$$

hence,

$$R - \int_a^b f = \int_{[a,b]} f$$

■

R Lebesgue integral indeed extends the Riemann integral in a theoretical perspective but knowing that Riemann and Lebesgue integrals do agree in values give us more ways to calculate Lebesgue integrals using Riemann integral tricks.



5. L^p Spaces

Throughout the course, f will be a measurable function defined on the measurable set E. First, $1 \leq p < \infty$ and notice that $|f|^p$ is again measurable by continuous composition on measurable function.

R Note that $p < \infty$, for $p = \infty$, we will have a separate definition for the L^∞ norm.

Definition 5.0.1 — L^p Seminorm.

$$\|f\|_{L^p(E)} = \left(\int_E |f|^p\right)^{1/p}$$

If E is clear, we write $\|f\|_p$.

R Note that we call this a seminorm since it is not true that $\|f\|_p = 0 \to f = 0$. We have $f = 0$ a.e., on E, then $\int_E |f|^p = 0$. To solve this problem to get a norm, we define an equivalence relationship on the measurable functions defined on E by specifying that

$$f \sim g, \text{ if } f - g = 0 \text{ a.e. on } E$$

Exercise 5.1

1. Please verify that if $\int_E |f|^p = 0$, then $f = 0$ a.e. on E and that if $f \sim g$, then $\|f\|_p = \|g\|_p$.
   Solution:
2. Show that if $\int_E |f|^p < \infty$, then $m\{x \in E : |f(x)| = \infty\} = 0$.

Definition 5.0.2

$$L^p(E) = \left\{\text{equivalence classes of measurable } f : E \to \mathbb{C} : \|f\|_{L^p(E)} < \infty\right\}$$

Again, we may drop E if it is clear.

Our end goal here is to prove that $L^p(E)$ is a Banach Space, i.e, a complete normed vector/linear space.

Proposition 5.0.1 $L^p(E)$ is a vector space.

Proof. Let $f, g \in L^p(E)$ and $\alpha \in \mathbb{C}$. Then, αf is measurable and $\|\alpha f\|_p = |\alpha|\|f\|_p < \infty$, so $\alpha f \in L^p$.
Clearly $f+g$ is measurable. By convex function property of $y = x^p$, for all $a, b \geq 0$,

$$\left(\frac{a+b}{2}\right)^p \leq \frac{a^p + b^p}{2} \iff (a+b)^p \leq 2^{p-1}(a^p + b^p)$$

Then,

$$\int |f+g|^p \leq \int (|f|+|g|)^p \leq 2^{p-1} \int |f|^p + |g|^p < \infty$$

Thus, $f+g \in L^p$. Therefore, $L^p(E)$ is a vector space. ■

R To have L^p to be normed vector space. We only need to show the triangle inequality works in L^p norm, $\|\cdot\|_p$.

Definition 5.0.3 — L^∞ Norm-Essential Supremum. Put

$$\|f\|_{L^\infty(E)} = \inf_{\substack{A \text{ measurable} \\ m(E \setminus A) = 0}} \{\sup |f(x)| : x \in A\}$$

we will write $\|f\|_\infty$ if E is clear.
If $E = \mathbb{R}$, we can write this as

$$\|f\|_{L^\infty(\mathbb{R})} = \inf_{m(A^c)=0} \{\sup |f(x)| : x \in A\}$$

R Note that by definition, we have $\|f\|_\infty \leq \sup_{x \in E} |f(x)|$. We shall ignore the function values on the measure zero sets.

■ **Example 5.1** Suppose $f = 0$ a.e. and $A = \{x : f(x) = 0\}$. Then, $\sup_{x \in A} |f(x)| = 0$ and $m(A^c) = 0$. Thus, $\|f\|_\infty = 0$. ■

Exercise 5.2 Show that if $f = g$ a.e., then $\|f\|_\infty = \|g\|_\infty$.
Solution: ■

Proposition 5.0.2 If f is continuous on $\mathbb{R}$, then $\|f\|_{L^\infty(\mathbb{R})} = \sup_x |f(x)|$

Proof. If $M < \sup |f|$, there exists x_0 and $m+1 > M$ such that $|f(x_0)| = M_1$. By continuity, there exists $\delta > 0$ such that $|f(x)| > M$ for all $x \in (x_0 - \delta, x_0 + \delta)$. Let A be any set with $m(A^c) = 0$.

Then, $(x_0-\delta, x_0+\delta)$ is not a subset of A^c (must have some intersection with A), then there exists $y \in (x_0-\delta, x_0+\delta) \cap A$. Then,

$$\sup_{x\in A} |f(x)| \geq |f(y)| > M$$

since A is arbitrary, we have $\|f\|_\infty \geq M$. Since this is true for all $M < \sup|f|$, we have $\|f\|_\infty \geq \sup|f|$. Inherently, we have $\|f\|_\infty \leq \sup|f|$. Thus, $\|f\|_\infty = \sup|f|$. ■

R The proof is almost identical to show the case when $E = [a,b]$ instead of $\mathbb{R}$.

Exercise 5.3 Show that

$$\|f\|_\infty = \inf\{\alpha \in \mathbb{R} : m\{x : |f(x)| > \alpha\} = 0\}$$

Solution: Delayed after assignment submission. ■

Definition 5.0.4 — L^∞ Space.

$$L^\infty(E) = \{\text{equivalence classes of measurable } f : E \to \mathbb{C}L\|f\|_\infty < \infty\}$$

R It is clear that every bounded measurable function is in the L^∞ space. In particular, the continuous function space over a compact set

$$\mathscr{C}([a,b]) \subseteq L^\infty([a,b])$$

But L^∞ can contain unbounded function, say

$$f(x) = \begin{cases} n & n \in \mathbb{N} \\ 0 & \text{otherwise} \end{cases}$$

is also in the L^∞ since $f = 0$ a.e.
We also call L^∞ the essentially bounded function space.

Proposition 5.0.3

$$\|f\|_\infty = 0 \implies f = 0 \text{ a.e.}$$

Proof. Using the result of Exercise 5.3, since $\|f\|_\infty = 0$, for each $n \in \mathbb{N}$, we have $m\{x : |f(x)| > \frac{1}{n}\} = 0$. However,

$$\{x : f(x) \neq 0\} = \bigcup_{n=1}^{\infty} \left\{x : |f(x)| > \frac{1}{n}\right\}$$

then, by $\sigma-$sub-additivity, we have

$$m\{x : f(x) \neq 0\} = 0$$

Thus, $f = 0$ a.e. ■

6. Hölder's Inequality

6.1 Hölder's Inequality

Hölder's Inequalitycan be considered to be a generalization of the famous Cauchy-Schwartz inequality.

Definition 6.1.1 — Conjugate. The pair p,q with $1 \leq p,q \leq \infty$ are called conjugate (or dual) indices if $\frac{1}{p} + \frac{1}{q} = 1$.

R Intuitively, we have $\frac{1}{\infty} = 0$. Some of the quick examples of conjugate pairs are

1. $(1,\infty)$
2. $(2,2)$
3. $(4,4/3)$

Theorem 13 — Hölder's Inequality. Suppose p,q are conjugate indices. If f,g are measurable functions, then

$$\int |fg| \leq \|f\|_p \|g\|_q = \left(\int |f|^p\right)^{1/p} \left(\int |g|^q\right)^{1/q}$$

In particular, if $f \in L^p$ and $g \in L^q$, then $fg \in L^1$ and $\|fg\|_1 \leq \|f\|_p \|g\|_q$

Proof.

1. When $p = 1, q = \infty$ (by symmetry), let $A = \{x : |g(x)| > \|g\|_\infty\}$, then $m(A) = 0$, by definition of the essential norm. Then,

$$\begin{aligned}\int |fg| &= \int_A |fg| + \int_{A^c} |fg| = \int_{A^c} |fg| \\ &\leq \|g\|_\infty \int_{A^c} |f| \leq \|g\|_\infty \|f\|_1\end{aligned}$$

2. If $1 < p,q < \infty$:

(a) **Step 1:** we prve the following inequality

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^1$$

for $a, b > 0$. Let $s, t \in \mathbb{R}$ such that $a = e^{s/p}$ and $b = e^{t/q}$. Then, by Jensen's inequality on convex function $\exp(\cdot)$,

$$ab = \exp\left(\frac{s}{p}\right)\exp\left(\frac{t}{q}\right) = \exp\left(\frac{1}{p}s + \frac{1}{q}t\right) \leq \frac{1}{p}e^s + \frac{1}{q}e^t = \frac{1}{p}a^p + \frac{1}{q}b^q$$

(b) **Step 2:** If $\|f\|_q = 0$ or $\|g\|_q = 0$, then $fg = 0$ a.e. and the result is clear. Similarly, if $\|f\|_p$ or $\|g\|_q = \infty$, the result is clear. Thus, we consider the other cases. Let $a = \frac{|f(x)|}{\|f\|_p}$ and $b = \frac{|g(x)|}{\|g\|_q}$ for each x. Thus, by Step 1,

$$\frac{|f||g|}{\|f\|_p\|g\|_q} \leq \frac{1}{p}\left(\frac{|f|}{\|f\|_p}\right)^p + \frac{1}{q}\left(\frac{|g|}{\|g\|_q}\right)^q$$

Note that $\|f\|_p$ is a constant, we see that

$$\int\left(\frac{|f|}{\|f\|_p}\right)^p = \frac{1}{\left(\|f\|_p\right)^p}\int|f|^p = 1$$

symmetrically,

$$\int\left(\frac{|g|}{\|g\|_q}\right)^q = 1$$

integrating both sides,

$$\int\frac{|f||g|}{\|f\|_p\|g\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1$$

then,

$$\|fg\|_1 = \int|fg| = \int|f||g| \leq \|f\|_p\|g\|_q$$

If $f \in L^p$ and $g \in L^q$, then this also shows $fg \in L^1$.

■

R

1. Proofs related to the L^p norms, usually the cases of $p = 1$ and/or $p = \infty$ need to be handled separately.
2. In the case of $p = q = 2$, we have

$$\left|\int f\overline{g}\right| \leq \int|f\overline{g}| \leq \|f\|_2\|g\|_2$$

which is a special case of the C-S inequality when using the inner product

$$\int f\overline{g} = \langle f, g\rangle$$

then, the Hölder's Inequalityis stating that

$$|\langle f, g\rangle|^2 \leq \sqrt{\langle f, f\rangle \cdot \langle g, g\rangle}$$

Exercise 6.1 Let $p > 1$,
1. prove that $\|f\|_{L^p[0,1]} \geq \|f\|_{L^1[0,1]}$
2. Show that $L^p[0,1]$ is a proper subset of $L^1[0,1]$
3. What can be said if $[0,1]$ is replaced by an arbitrary compact interval $[a,b]$?

Theorem 14 Let $1 \leq p < \infty$ and suppose q is a conjugate index. For any $f \in L^p$,

$$\|f\|_p = \sup\left\{\left|\int fg\right| : \|g\|_q \leq 1\right\}$$

Proof. Delayed to PMATH451. ■

R This is almost like a converse of the Hölder's Inequalityinequality.

6.2 Minkowski's Inequality

Minkowski's inequality is the triangle inequality for the L^p spaces.

Theorem 15 — Minkowski Inequality. For f, g measurable,

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p, \forall 1 \leq p \leq \infty$$

Proof.
1. **Case:** $p = \infty$ Delayed after assignment submitted.
2. **Case:** $p = 1$.

$$\|f+g\|_1 = \int |f+g| \leq \int |f| + |g| = \int |f| + \int |g| = \|f\|_1 + \|g\|_1$$

3. **Case:** $1 < p < \infty$. We may assume $f, g \in L^p$ otherwise the right hand side is infinite and the inequality is clearly true. Then, $f + g \in L^p$ by its vector space property. Likewise, assume $\|f+g\|_p > 0$. Let q be the conjugate index to p. Then,

$$\begin{cases} 1 + \frac{p}{q} = p \\ p + q = pq \\ p = q(p-1) \end{cases}$$

Then,

$$\begin{aligned}\left(\|f+g\|_p\right)^p &= \int |f+g|^p = \int |f+g|^{p-1}|f+g| \\ &\leq \int |f+g|^{p-1}(|f|+|g|)\end{aligned}$$

we have another inequality by Hölder's Inequality,

$$\int |f+g|^{p-1}|f| \leq \left\||f+g|^{p-1}\right\|_q \|f\|_p \tag{6.2.1}$$

then,

$$\left(\left\| |f+g|^{p-1} \right\|_q\right)^q = \int |f+g|^{(p-1)q} = \int |f+g|^p = \|f+g\|_p^p < \infty$$

then,

$$\int |f+g|^{p-1}|f| \leq \|f+g\|_p^{p/q}\|f\|_p$$

by symmetry,

$$\int |f+g|^{p-1}|g| \leq \|f+g\|_p^{p/q}\|g\|_p$$

Then,

$$\|f+g\|_p^p \leq \|f+g\|_p^{p/q}\left(\|f\|_p + \|g\|_p\right)$$

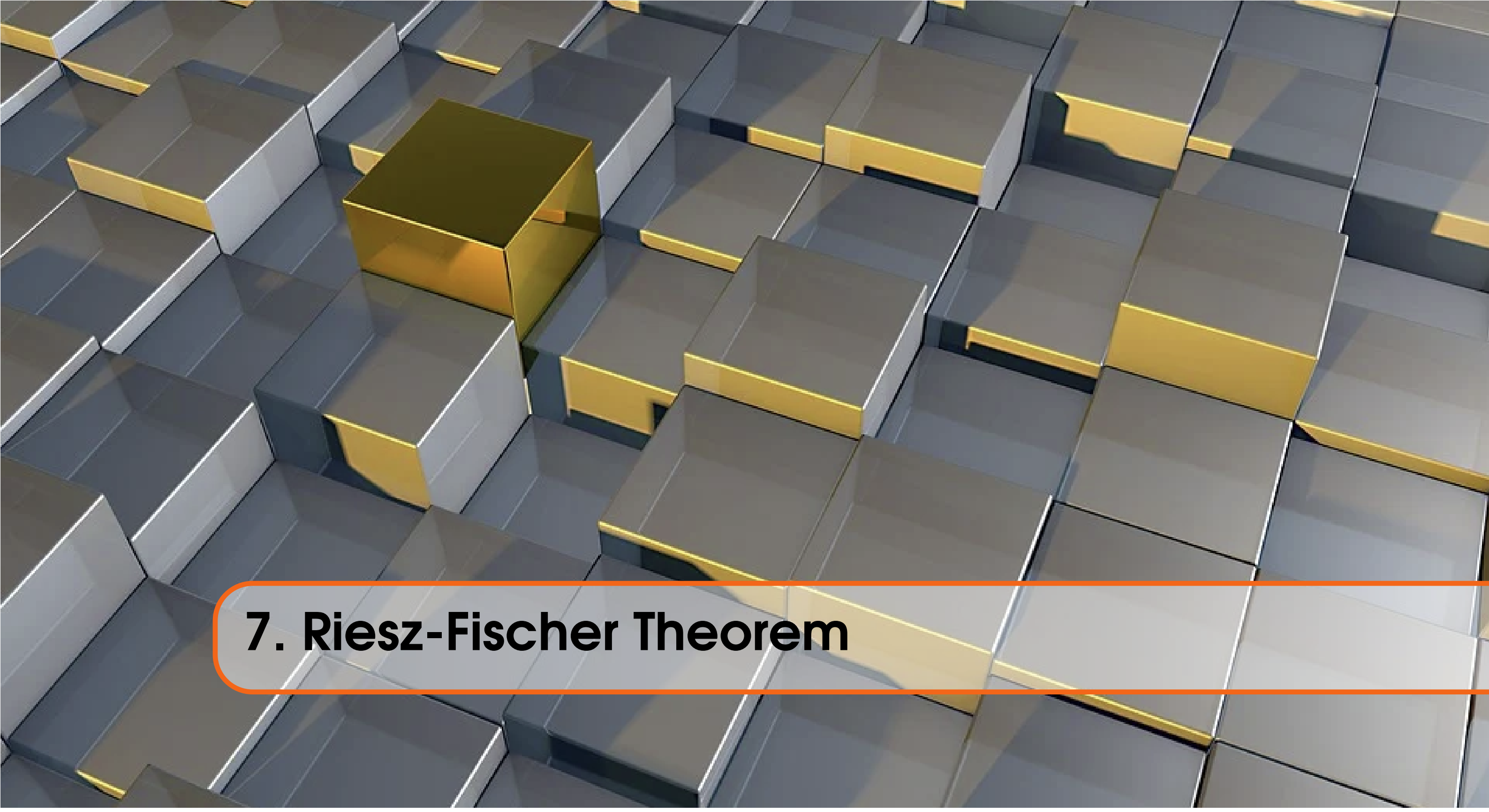
$$\|f+g\|_p^{p-p/q} = \|f+g\|_p \leq \|f\|_p + \|g\|_p$$

and we are done. ■

Corollary 6.2.1 For $1 \leq p \leq \infty$, L^p is a normed linear space, which is also a metric space using the induced metric

$$d(f,g) = \|f-g\|_p$$

Proof. A PMATH351 result. ■



7. Riesz-Fischer Theorem

Our original goal was to show every L^p space is a Banach space. We have shown it is a normed linear space, it remains to show its completeness.

Theorem 16 — Riesz-Fischer Theorem. L^p is a Banach space for any $1 \le p \le \infty$.

Proof. 1. **Case $p = \infty$:** Let (f_n) be a Cauchy sequence in $L^\infty$and for $k, n, m \in \mathbb{N}$ set

$$A_k = \{x : |f_k(x)| > \|f_k\|_\infty\}$$
$$B_{n,m}\{x : |f_n(x) - f_m(x)| > \|f_n - f_m\|_\infty\}$$

these sets are of measure 0, then the countable union of them, E, is for sure of measure 0. We have

$$\sup_{x\in E^c} |f_n(x) - f_m(x)| \le \|f_n - f_m\|_\infty \to 0, \text{ as } n, m \to \infty$$

Thus, the seqeunce (f_n) is uniformly Cauchy on E^c. Uniform Cauchy seqeunces of real/complex-valued functions always converge uniformly, meaning there is some f defined on E^c such that

$$\sup_{E^c} |f_n(x) - f(x)| \to 0$$

We define $f(x) = 0$ on E. Since $m(E) = 0$, we have

$$\|f_n - f\|_\infty \le \sup_{x\in E^c} |f_n(x) - f(x)| \to 0$$

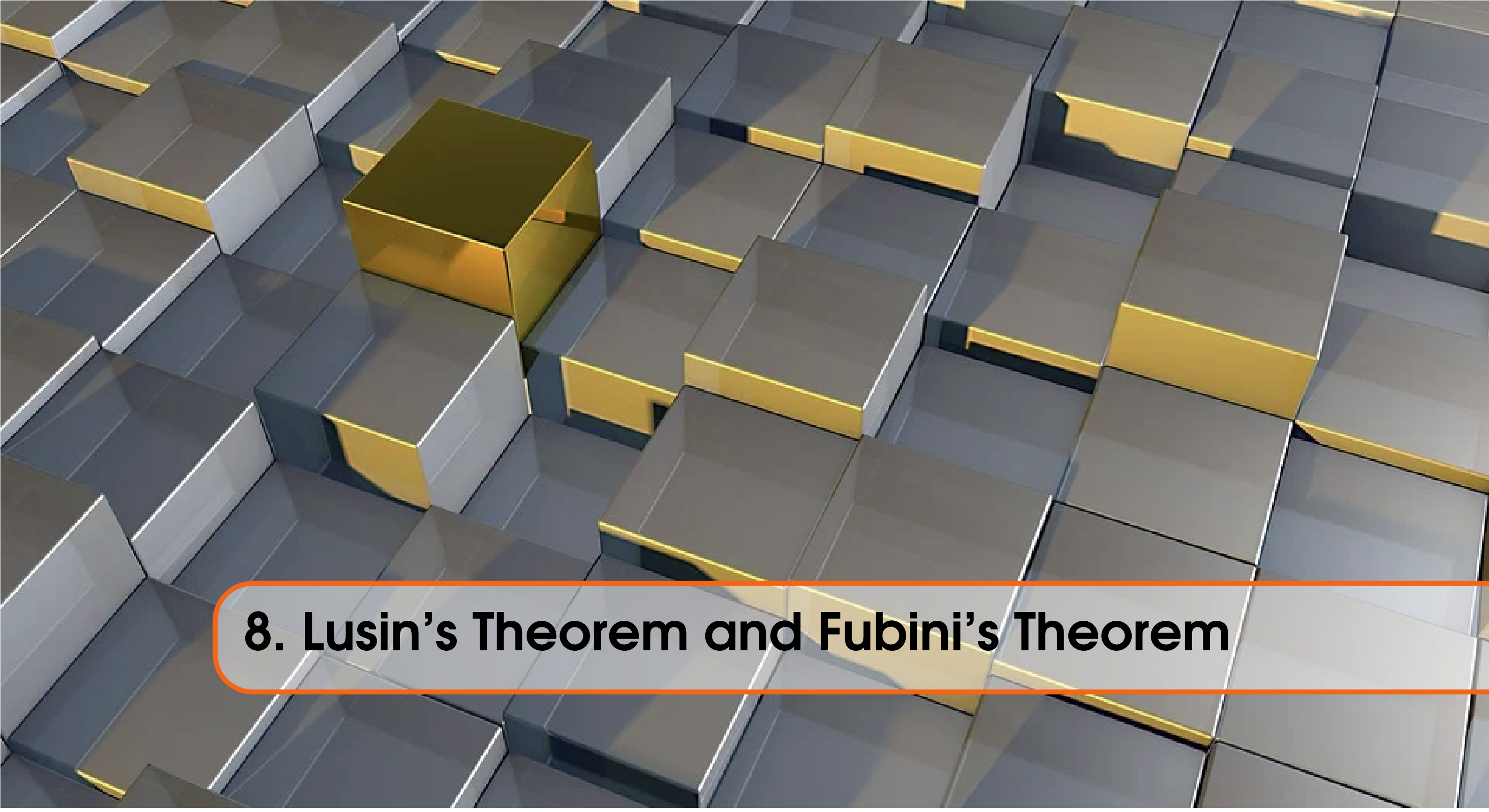
Thus, $f_n \to f$ in L^∞. It remains to show that the limit $f \in L^\infty$.
We can choose N such that $\sup_{x\in E^c} |f_N(x) - f(x)| \le 1$. Since $A_N \subseteq E$, the by triangle inequality,

$$\sup_{x\in E^c} |f(x)| \le \sup_{x\in E^c} (|f_N(x) - f(x)| + |f_N(x)|) \le 1 + \|f_N\|_\infty < \infty$$

2. **Case $p < \infty$:** let (f_n) be a Cauchy sequence in L^p. Using the Cauchy property, we choose a subsequence (f_{n_i}) such that
$$\left\|f_{n_{i+1}} - f_{n_i}\right\|_p < 2^{-i}$$
set
$$g_k = \sum_{i=1}^{k} \left|f_{n_{i+1}} - f_{n_i}\right|, g = \sum_{i=1}^{\infty} \left|f_{n_{i+1}} - f_{n_i}\right|$$
it might be possible that $g(x)$ to be $+\infty$ but this is okay. By Minkowski inequality,
$$\|g_k\|_p \leq \sum_{i=1}^{k} \left\|f_{n_{i+1}-f_{n_i}}\right\|_p \leq \sum_{i=1}^{k} 2^{-i} \leq i, \forall k$$
Then, $g_k \uparrow g$ and $g)k^p \uparrow g^p$ pointwise. By MCT,
$$\|g\|_p^p = \int g^p = \lim_k \int g_k^p = \lim_k \|g_k\|_p^p \leq 1$$
thus, $g \in L^p$. Thus, $g(x) < \infty$ a.e. Then, almost everywhere we have
$$\sum_{i=1}^{\infty} \left|f_{n_{i+1}}(x) - f_{n_i}(x)\right|$$
is absolutely convergent, which implies $\sum_{i=1}^{\infty} f_{n_{i+1}}(x) - f_{n_i}(x)$ converges by MATH148. Define the function f on the A^c (the almost everywhere set) as the pointwise limit
$$f(x) = \begin{cases} \lim_k \left(f_{n_1}(x) + \sum_{i=1}^{k} \left(f_{n_{i+1}}(x) - f_{n_i}(x)\right)\right) = \lim_k f_{n_k}(x) & x \in A^c \\ 0 & x \in A \end{cases}$$
f is a measurable function (exercise, we can write $f = f\chi_{A^c} + f\chi_A$). For all $x \in A^c$,
$$|f(x)| \leq |f_{n_1}(x)| + \sum_{i=1}^{\infty} \left|f_{n_{i+1}}(x) - f_{n_1}(x)\right| = |f_{n_1}(x)| + g(x) \Longrightarrow f \in L^p$$
f is our candidate for the limit. We have seen taht $f_{n_k}(x) \to f(x)$ for all $x \in A^c$ and that means $|f_{n_k} - f| \to 0$ a.e. Also, on A^c,
$$|f_{n_k} - f| \leq \sum_{i=k}^{\infty} \left|f_{n_{i+1}} - f_{n_i}\right| \geq g$$
so $|f_{n_k} - f|^p \leq g^p \in L^1(A^c)$ since $g \in L^p$. By DCT,
$$\lim_k \int_{A^c} |f_{n_k} - f|^p = \int_{A^c} \lim_k |f_{n_k} - f|^p = 0$$
since $m(A) = 0$, we have $\|f_{n_k} - f\|_p \to 0$ as required.

■

Corollary 7.0.1 — Corollary of the Proof. If $f_n \to f$ in L^p for $1 \leq p \leq \infty$, then there exists a subsequence f_{n_k} that converges to f pointwise almost everywhere. In the case of $p = \infty$, the whole seqeunce $f_n \to f$ pointwise almost everywhere.



8. Lusin's Theorem and Fubini's Theorem

8.1 Statement of Lusin's Theorem

We have seen that every bounded function is essentially bounded and thus $C[a,b] \subseteq L^\infty[a,b]$. Moreover, $L^\infty[a,b] \subseteq L^p[a,b]$ for all $p < \infty$ since

$$\left(\int_{[a,b]} |f|^p\right)^{1/p} \leq \left(\|f\|_\infty^p \int_{[a,b]} 1\right)^{1/p} = \|f\|_\infty (b-a)^{1/p}$$

Thus, $C[a,b] \subseteq L^p[a,b]$ for all $p \geq 1$. There are discontinuous functions in L^p, hence it is natural to ask if $C[a,b]$ is a 'large' subset of L^p. Lusin's theorem addresses this question.

R Note that $\frac{1}{\sqrt{x}} \in L^1[0,1]$ but not in $L^\infty[0,1]$ (for any equivalence class of it)

Definition 8.1.1 — Dense in L^p. To be dense in L^p for $p < \infty$ means that for every $f \in L^p[a,b]$ and for every $\varepsilon > 0$, there exists $g \in C[a,b]$ such that

$$\|f - g\|_{L^p[a,b]} < \varepsilon$$

or equivalently,

$$\overline{C[a,b]} = L^p[a,b]$$

or equivalently, there exists $(g_n) \in C[a,b]$ such that $\|g_n - f\|_p \xrightarrow{n\to\infty} 0$.

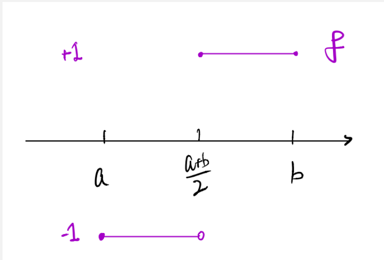
Theorem 17 — Lusin's Theorem. The continuous functions are dense in $L^p[a,b]$ for $1 \leq p < \infty$, but are NOT dense in $L^\infty[a,b]$.

Proof. **Proof that $C[a,b]$ is not dense in $L^\infty[a,b]$**

Proof. Consider the function

$$f(x) = \begin{cases} -1 & \left[a, \frac{a+b}{2}\right] \\ +1 & \left(\frac{a+b}{2}, b\right] \end{cases}$$

such jump discontinuity will result in $f = g$ a.e., then g is also not continuous. Thus, the equivalence containing f in L^∞ does not contain continuous function on $[a,b]$. Why is that? Consider the graph below,



say $f = g$ a.e. for some continuous g , assume $f = g$ on A^c where $m(A) = 0$. Note that

$$\left(\frac{a+b}{2} - \frac{1}{n}, c\right) \cap A^c \neq \emptyset, \forall n$$

since $m(A) = 0$. Let $x_n \in \left(\frac{a+b}{2} - \frac{1}{n}, c\right) \cap A^c$. Similarly, for all n, there exists $y_n \in \left(c, \frac{a+b}{2} + \frac{1}{n}\right) \cap A^c$. Then, by sequential continuity,

$$\begin{cases} -1 = f(x_n) = g(x_n) \xrightarrow{n\to\infty} g\left(\frac{a+b}{2}\right) \\ +1 = f(y_n) = g(y_n) \xrightarrow{n\to\infty} g\left(\frac{a+b}{2}\right) \end{cases}$$

this yields a contradiction.

Back to the main proof, suppose that the sequence (f_n) belongs to $C[a,b]$ converges to f in the L^∞ norm. Then (f_n) is Cauchy in the L^∞ norm. But as the function f_n are continuous, this means (f_n) is Cauchy in the sup/uniform norm on $C[a,b]$. Bit $C[a,b]$ with the sup norm is complete, hence the sequence (f_n) must converges to some continuous function g in the sup norm. Since the sup norm and the L^∞ norm coincide on continuous functions, that means,

$$\|f_n - g\|_{L^\infty[a,b]} = \sup_{x\in[a,b]} |f_n(x) - g(x)| \to 0$$

then, g is the limit of (f_n) in the L^∞ norm. But limits in metric spaces are unique, hence as elements of L^∞ it must be taht f and g coincide or say $f = g$ a.e. But we have shown that the equivalence class containing f does not contain any continuous function. This yields a contradiction. ■

Proof that $C[a,b]$ is dense in $L^p[a,b]$ for $p < \infty$

Lemma 8.2 Suppose $m(E) < \infty$. Then, for every $\delta > 0$, there is a finite union of open itnervals U suhc that $m(U \backslash E) + m(E \backslash U) < \delta$.

Proof. Proof. Choose $\{I_n\}$ open intervals whose union contains E and for which

$$\sum_n l(I_n) < m(E) + \frac{\delta}{2} < \infty$$

Thus, $m(\cup I_n \backslash E) < \frac{\delta}{2}$. Choose N such that $\sum_{n=N+1}^{\infty} l(I_n) < \frac{\delta}{2}$ and put $U = \cup_{n=1}^{N} I_n$. We check

$$U \backslash E \subseteq \left(\bigcup_{n=1}^{\infty} I_n \right) \backslash E$$

$$E \backslash U \subseteq \left(\bigcup_{n=1}^{\infty} I_n \right) \backslash U = \bigcup_{n=N+1}^{\infty} I_n$$

then, by monotinicity, we have

$$m(U \backslash E) + m(E \backslash U) < \frac{\delta}{2} + \frac{\delta}{2} = \delta$$

■

Lemma 8.3 Let $f : [a,b] \to \mathbb{R}$ be measurable. Fix $\varepsilon > 0$, there exists a continuous function h such that

$$m\{x \in [a,b] : |f(x) - h(x)| \geq \varepsilon\} < \varepsilon$$

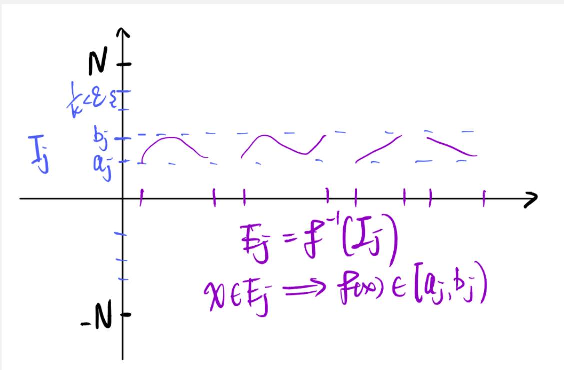
moreover, if $m \leq f \leq M$, then we can choose with $m \leq h \leq M$.

Proof.

1. **Step 1:** Show there exists N such that $m\{x \in [a,b] : |f(x)| \geq N\} < \frac{\varepsilon}{3}$.
   For each $n \in \mathbb{N}$, let $A_n = \{x \in [a,b] : |f(x)| \geq n\}$. The sets A_n are decreasing and $m(A_1) \leq b - a < \infty$. Thus, by downward continuity of measure, $m(\cap_n A_n) = \lim_n m(A_n)$, but

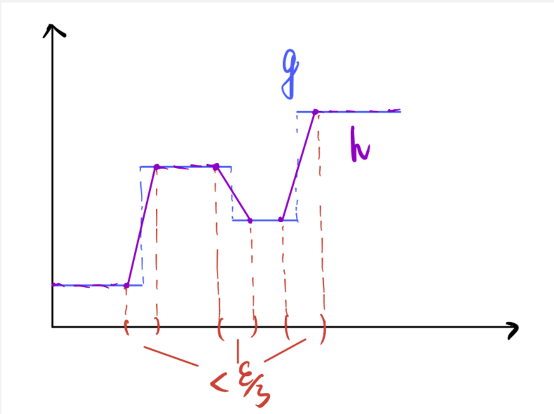
   $$\bigcap_{n=1}^{\infty} A_n = \{x \in [a,b] : |f(x)| \geq n, \forall n \in \mathbb{N}\} = \emptyset$$

   thus, $m(A_n) \to 0$.
2. **Step 2:** show that for each $\varepsilon > 0$ and $N \in \mathbb{N}$, there exists a simple function ϕ such that $|f(x) - \phi(x)| < \varepsilon$ if $|f(x)| < N$. Furthermore, if $m \leq f \leq M$m then we can choose ϕ with $m \leq \phi \leq M$.
   Pick $k \in \mathbb{N}$ with $\frac{1}{k} < \varepsilon$. Partition $[-N,N)$ into subintervals of width $\frac{1}{k}$ denoted by $I_j = [a_j, b_j)$ for $j = 1, 2, \cdots, 2Nk$. Let E_j be the measurable set $f^{-1}(I_j)$ and define $\phi(x) = a_j$ for $x \in E_j$. Then, $|f(x) - \phi(x)| \leq \frac{1}{k} < \varepsilon$. (Consider the picture below)



Note that this has already proved the first part of the step 2. For the second part, we just need to repartition $[m,M]$ in a similar fashion.

3. **Step 3:** Show that given any simple function ϕ defined on $[a,b]$, there is a step function g on $[a,b]$ such that $g(x) = \phi(x)$ except on a set of measure $< \frac{\varepsilon}{3}$. Again, if $m \le \phi \le M$, then we can choose g with $m \le g \le M$.
   Assume $\phi = \sum_{k=1}^{K} a_k \chi_{E_k}$ where $E_k \subseteq [a,b]$ are measurable sets. For each k, by Lemma 8.2, witht he sets E_k and $\delta = \frac{\varepsilon}{3K}$, to obtain sets U_k, each of which is a finite union of open intervals. Put $g = \sum_{k=1}^{K} a_k \chi_{U_k}$. As the U_k are finite unions of intervals, g is a step function. By construction $\phi = g$ except on $\cup_{k=1}^{K}(U_k \backslash E_k) \cup (E_k \backslash U_k)$. Then, the claim follows from this.
4. **Step 4:** show that given any step function g and $\varepsilon > 0$, there is a continuous function h such that $g = h$ except on a set of measure $< \frac{\varepsilon}{3}$. Moreover, if $m \le g \le M$, then we can choose h with $m \le h \le M$. See the picture below: basically, we just need to connect the jump discontinuity (fintiely many of them) with a line segment by controlling the domain of each segment.



5. **Step 5: complete the lemma proof** we pick N as in step 1. Then, choose ϕ, g, h and all the steps above showed that

$$|f - h| \le \varepsilon$$

   except on the union of sets $\{x \in [a,b] : \phi(x) \ne g(x)\}$, $\{x \in [a,b] : g(x) \ne h(x)\}$, and $\{x \in [a,b] : |f(x)| \ge N\}$, but the measure of these sets summed up to ε.

■

Finally, back to the Lusin's Theorem! Let $f \in L^p[a,b]$. Then, $|f| < \infty$ a.e. (otherwise, the L^p norm will blow up to infinity when integrating on such a positive set). Then, we can choose another function from the same equivalence class to not have $\pm\infty$ values. Say $f : [a,b] \to \mathbb{C}$. Then, $\operatorname{Re} f, \operatorname{Im} f : [a,b] \to \mathbb{R}$ belong to L^p. It is enough to approximate each of $\operatorname{Re} f, \operatorname{Im} f$ and invoke triangle inequality to finish it off. Thus, WLOG, say $f : [a,b] \to \mathbb{R}$.

1. f **is bounded:** First, suppose f is bounded, say $-N \le f(x) \le N$ for all x. Let $\varepsilon > 0$ and put

$$\delta = \min\left(\frac{\varepsilon}{(2N)^p}, \frac{\varepsilon}{b-a}, 1\right) > 0$$

   by Lemma 8.3, we can obtain a continuous function h with $\delta = \varepsilon$. Let $A = \{x : |f - h| \ge \delta\}$, then

$$m(A) < \delta \text{ and } -N \le h \le N$$

we now check

$$\|f-h\|_p^p = \int_{[a,b]} |f-h|^p = \int_A |f-h|^p + \int_{A^c} |f-h|^p$$
$$\leq \int_A (|f|+|h|)^p + \delta^p \int_{A^c} 1$$
$$\leq (2N)^p m(A) + \delta^p (b-a)$$
$$\leq \delta((2N)^p + b - a) \leq 2\varepsilon$$

2. **Suppose f is not bounded:** put

$$f_n(x) = \begin{cases} f(x) & |f(x)| \leq n \\ 0 & \text{otherwise} \end{cases}$$

for $\varepsilon > 0$ fixed, note that $f_n \to f$ pointwise a.e. and $|f_n - f|^p \leq |f|^p \in L^1$ since $f \in L^p$. Hence by the DCT,

$$\int_{[a,b]} |f_n - f|^p = 0$$

choose N such that $\|f_N - f\|_p < \varepsilon$. As $f_N \in L^p$ is bounded, by the bounded case, we have a continuous h such that $\|f_N - h\|_p < \varepsilon$. By the Minkowski's inequality gives

$$\|f-h\|_p \leq \|f - f_N\|_p + \|f_N - h\|_p < 2\varepsilon$$

This completes the proof of the Lusin's Theorem when $p < \infty$. ■

■

Corollary 8.3.1 The polynomials are dense in $L^p[0,1]$ for $1 \leq p < \infty$. (Note that this is in the L^p norm sense)

Proof. Recall that

$$\|f\|_{L^p[0,1]} = \left(\int_{[0,1]} |f|^p\right)^{1/p} \leq \|f\|_{L^\infty[0,1]} \left(\int_{[0,1]} 1\right)^{1/p} = \|f\|_\infty$$

Fix $\varepsilon > 0$ and let $f \in L^p$. Choose a continuous function h such that $\|f-h\|_p < \varepsilon$. Now apply the Stone Weierstrass theorem to obtain a polynomial P such that

$$\sup\{|P(x) - h(x)| : x \in [0,1]\} < \varepsilon$$

since $\sup_x |P(x) - h(x)| = \|P-h\|_\infty \geq \|P-h\|_p$, we have

$$\|f-P\|p \leq \|f-h\|_p + \|h-P\|_p \leq \varepsilon + \sup_x |P(x) - h(x)| \leq 2\varepsilon$$

Thus, we have the polynomials are dense. ■

Exercise 8.1 Check that $C[a,b]$ is a closed subset of $L^\infty[a,b]$.
Solution: ■

Exercise 8.2
1. Prove that $L^p[0,1]$ is separable (has a countable dense subset) for $1 \leq p < \infty$).
   Solution:
2. Prove that $L^\infty[0,1]$ is not separable.
   Solution:

■

R This theorem is a formal version of what is sometimes referred to as the second of Littlewood's three principles. These principles are:
1. Every [measurable] set is 'nearly' a finite union intervals.
2. Every [measurable] function is 'nearly' continuous
3. Every pointwise convergent sequence of [measurable] functions is 'nearly' uniformly convergent

Theorem 18 — Egoroff's Theorem. Assume $m(E) < \infty$. Let (f_n) be a sequence of measurable function on E that converges pointwise on E to the real-valued function f. For each $\varepsilon > 0$, there is a closed set $F \subseteq E$ for which $(f_n) \to f$ uniformly on F and $m(E \backslash F) > \varepsilon$.

8.4 Fubini's Theorem

We have seen Fubini's theorem in MATH247/PMATH365 for the Riemann integral. This theorem generalizes to the Lebesgue integral. By a Borel set in $\mathbb{R}^2$, we mean any set in the $\sigma-$algebra of subsets of $\mathbb{R}^2$ generated by the open sets in $\mathbb{R}^2$.

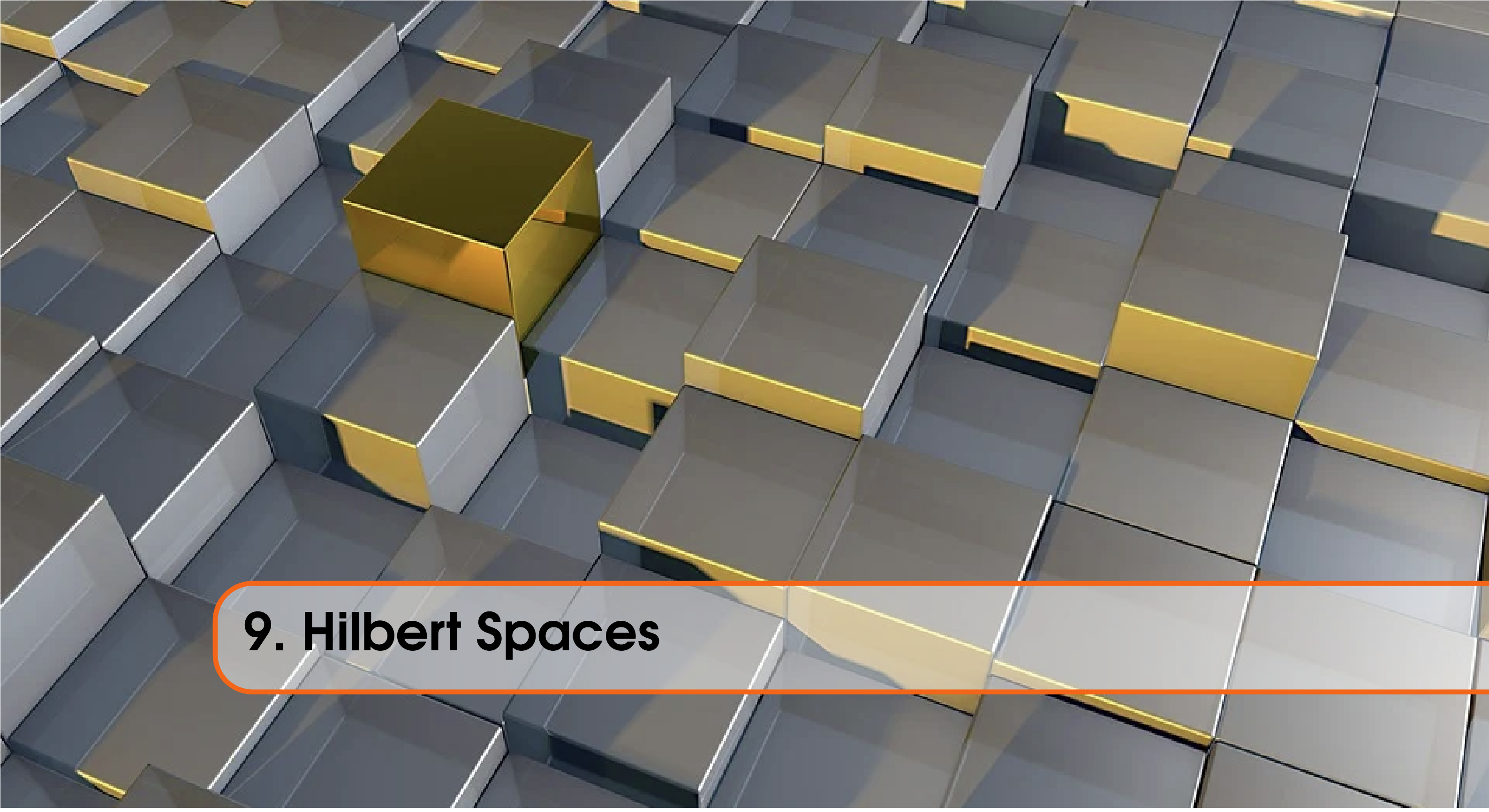
Theorem 19 — Fubini's Theorem. Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ and $f^{-1}(U)$ is a Borel set in $\mathbb{R}^2$ for all open sets $U \subseteq \mathbb{R}$. Assume that

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x,y)| dx \right) dy < \infty$$

then,

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x,y) dx \right) dy = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x,y) dy \right) dx$$

R This theorem states that under modest assumptions, we can interchange the order of integration of a function of two variables. It will be proven in PMATH451 in greater generality.



9. Hilbert Spaces

9.1 Inner Products and Orthogonality

A Hilbert space is a special example of a Banach space where the norm arises from an inner product.

Definition 9.1.1 — Inner Product Space (IPS). Let V be a complex vector space. An inner product on V is a map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ with the following properties:

1. $\langle \alpha f, g \rangle = \alpha \langle f, g \rangle, \forall \alpha \in \mathbb{C}, f, g \in V$
2. $\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle, \forall f, g, h \in V$
3. $\langle f, g \rangle = \langle g, f \rangle, \forall f, g \in V$
4. $\langle f, f \rangle \geq 0$ and $\langle f, f \rangle = 0$ if and only if $f = 0$.

Note that properties 1, 2, 3 imply

1. $\langle f, \beta g \rangle = \overline{\langle \beta g, f \rangle} = \overline{\beta} \langle f, g \rangle, \forall \beta \in \mathbb{C}$ and $f, g \in V$.
2. $\langle f, g + h \rangle = \overline{\langle g + h, f \rangle} = \langle f, g \rangle + \langle f, h \rangle, \forall f, g, h \in V$

■ **Example 9.1** $\mathbb{C}^n$ with $\langle x, y \rangle = \sum_{i=1}^{n} x_i \overline{y_i}$ ■

■ **Example 9.2** Let $l^2 = \left\{ (x_n)_{n=1}^{\infty} : \left(\sum_n |x_n|^2 \right)^{1/2} < \infty \right\}$ and define $\langle x, y \rangle = \sum_i x_i \overline{y_i}$. This is an important example for us. ■

■ **Example 9.3** $C[0,1]$ with $\langle f, g \rangle = \int_0^1$. As f, g are continuous, this integral can be taken to be either the Riemann or Lebesgue integral. ■

■ **Example 9.4** The most important example for our course will be $L^2[0,1]$ with $\langle f, g \rangle = \int_0^1 f\overline{g}$. Here the integral is the Lebesgue intergral. As $f, g \in L^2$, Hölder's Inequalityimplies that fg is integrabl since

$$\left| \int_0^1 f\overline{g} \right| \leq \int_0^1 |f\overline{g}| \leq \|f\|_2 \|g\|_2 < \infty$$

Thus, $\int_0^1 f\overline{g}$ is well defined. ■

Exercise 9.1 Prove Example 9.4 is actually an IPS. Especially,

$$\int \overline{g} = \overline{\int g}$$

■

R Every IPS gives you a norm by

$$\|f\| = (\langle f, f \rangle)^{1/2}$$

Exercise 9.2 Prove $\|f\|$ is actually a norm. ■

R We used to derive C-S inequality from Hölder's Inequality, but in fact, we can also derive it from IPS.

Definition 9.1.2 — Hilbert Space. An IPS that is complete. It is a Banach space where the norm comes from an inner product.

■ **Example 9.5** $\mathbb{C}^n$, l^2 (proved in PMATH351), and $L^2[0,1]$ (by Riesz-Fischer Theorem) are all Hilbert spaces. $C[0,1]$ with the norm coming from the norm coming from the IP is not a Hilbert space. In fact, Lusin's Theorem implies that $C[0,1]$ is a dense subspace of $L^2[0,1]$ with this norm. ■

Lemma 9.2 Fix $z \in \mathscr{H}$, the map $T : \mathscr{H} \to \mathbb{C}$ given by $T(y) = \langle y, z \rangle$ is a continuous linear map.

Proof.

1. **Linearity:**

$$T(\alpha y + w) = \langle \alpha y + w, z \rangle = \alpha \langle y, z \rangle + \langle w, z \rangle = \alpha T(y) + T(z)$$

2. **Continuity:** let $y_n \to y \in \mathscr{H}$. Then,

$$|T(y_n) - T(y)| = |\langle y_n - y, z \rangle| \underset{\text{C-S Inequality}}{\leq} \|y_n - y\| \|z\| \to 0$$

since $\|y_n - y\| \to 0$

■

R We want to emphasize that Hilbert space has certain taste of 'geometry' like the Euclidean geometry.

Definition 9.2.1 — Orthogonal. Let $x, y \in \mathscr{H}$ by orthogonal denoted by $x \perp y$ if $\langle x, y \rangle = 0$.

■ **Example 9.6**

1. $(0,1)$ and $(1,0)$ in $\mathbb{C}^2$ are orthogonal
2. χ_E and χ_{E^c} are orthogonal in $L^2[0,1]$. More generally, any pair $f, g \in L^2$ with $f\overline{g} = 0$ a.e. are orthogonal in L^2
3. In any Hilbert space, $x \perp x \iff x = 0$

■

Definition 9.2.2 — Orthogonal Set and Orthonormal Set. We say that $S \subseteq H$ is orthogonal, if $x \perp y$ for all $x, y \in S$ for any $x \neq y$. If, in addition, $\|x\| = 1$ for every $x \in S$, we call S othornormal.

■ **Example 9.7** Suppose $\mathscr{H} = l^2$. Put $e_n = (0, 0, \cdots, 1, 0, \cdots$ where 1 is in the $n-$th component and let $S = \{e_n : n \in \mathbb{N}\}$. This is an orthonormal set, the standard coordiante vectors for l^2. ■

Theorem 20 — Pythagorean Theorem. If $x \perp y$, then $\|x+y\|^2 = \|x\|^2 + \|y\|^2$. More generally, if $\{x_n\}_{n=1}^N$ is orthogonal, then

$$\left\|\sum_{n=1}^N x_n\right\|^2 = \sum_{n=1}^N \|x_n\|^2$$

Proof. Quite immediate from the definition and use it inductively to get the second part. ■

R Suppose $\{x_n\}_{n=1}^N$ are orthogonal and non-zero vectors and suppose $\sum_{n=1}^N \alpha_i x_i = 0$. Then,

$$0 = \left\|\sum_{n=1}^N \alpha_i x_i\right\|^2 = \sum_{i=1}^N \|\alpha_i x_i\|^2$$

and therefore, $|\alpha_i|\|x_i\| = 0$ implies $a_i = 0, \forall i$

9.3 Bessel's Inequality

Proposition 9.3.1 1. Let $\{x_k\}_k \subseteq \mathscr{H}$ be orthogonal vectors and for each N, let $S_N = \sum_{k=1}^N x_k \in \mathscr{H}$. Then $(S_N)_N$ convergs if and only if $\sum_{k=1}^\infty \|x_k\|^2 < \infty$

2. Let $\{e_k\}_k \subseteq \mathscr{H}$ be anorthonormal set and let $\beta_k \in \mathbb{C}$ with $\sum_k |\beta_k|^2 < \infty$. Then, there is some $x \in \mathscr{H}$ such that $\langle x, e_k\rangle = \beta_k$ for all k and $\|x\| = \|(\beta_k)\|_{l^2}$

Proof. 1. Let $N > M$ and consider $S_N - S_M = \sum_{k=M+1}^N x_k$. The Pythagorean theorem implies

$$\|S_N - S_M\|^2 = \left\|\sum_{k=M+1}^N x_k\right\|^2 = \sum_{k=M+1}^N \|x_k\|^2 = \left|\sum_{k=1}^N \|x_i\|^2 - \sum_{k=1}^M \|x_k\|^2\right|$$

we can see that the last quantity is a Cauchy difference of $\left(\sum_{k=1}^N \|x_k\|^2\right)_N$ which is in l^2. We know that $\mathscr{H}, l^2$ are both Hilbert spaces. Then, $(S_N)_N$ converges if and only if $(S_N)_N$ is Cauchy if and only if $\left(\sum_{k=1}^N \|x_k\|^2\right)_N$ is Cauchy if and only if $\sum_{k=1}^\infty \|x_k\|^2 < \infty$.

2. By orthogonality, we have

$$\sum \|\beta_k e_k\|^2 = \sum |\beta_k|^2 < \infty$$

then, $\sum \beta_k e_k$ converges, say to $x \in \mathscr{H}$. Then,

$$\left\|x - \sum_{k=1}^N \beta_k e_k\right\| \to 0$$

Thus,

$$\|x\|^2 = \lim_N \left\|\sum_{k=1}^N \beta_k e_k\right\|^2 = \lim_N \sum_{k=1}^N |\beta_k|^2 = \|(\beta_k)\|_{l^2}^2$$

we check that $\langle x, e_k\rangle = \beta_k$. By Lemma 9.2, we define $T_j(y) = \langle y, e_j\rangle$. Let $X_n = \sum_{k=1}^N \beta_k e_k$. Since $x_N \to x$ in $\mathscr{H}$. We have

$$\langle x_N, e_j\rangle \to \langle x, e_j\rangle$$

for each e_j. But we know that

$$\beta_j = \sum_{k=1}^N \langle \beta_k e_k, e_j\rangle = \langle x_N, e_j\rangle, (j \le N)$$

■

Theorem 21 — Bessel's Inequality. If $\{e_k\}_k$ is an orthonormal set and $x \in \mathscr{H}$, then

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2 \leq \|x\|^2 < \infty$$

In particular, the sequence $(\langle x, e_k \rangle)_k \in l^2$.

Proof. Let $x_N = \sum_{j=1}^{N} \langle x, e_j \rangle e_j$. If $N \geq k$, then $\langle x - x_N, e_k \rangle = \langle x, e_k \rangle - \langle x_N, e_k \rangle = 0$, since we have

$$\langle x_N, e_k \rangle = \sum_{j=1}^{N} \langle \langle x, e_j \rangle e_j, e_k \rangle = \sum_{j=1}^{N} \langle x, e_j \rangle \langle e_j, e_k \rangle = \langle x, e_k \rangle$$

Thus, $x - x_N \perp e_k$ for all $k \leq N$. In particular, $x - x_N \perp x_N$, so

$$\|x\|^2 = \|x - x_N\|^2 + \|x_N\|^2 \geq \|x_N\|^2 = \sum_{k=1}^{N} |\langle x, e_k \rangle|^2$$

let $N \to \infty$, we have the desired inequality. ■

9.4 Bases for Hilbert Spaces

R The notion of linearly independent spanning set of a basis will be considered as **algebraic basis**, which is not suitable for general Hilbert space. For any finite dimensional IPS, we can always find a orthonormal algebraic basis using the Gram Schmidt. But...

Proposition 9.4.1 An orthonormal set never spans an infinite dimensional Hilbert space.

Proof. Suppose $\{e_n\}_n$ is any orthonormal set. Then, $x = \sum_{n=1}^{\infty} \frac{1}{n} e_n \in \mathscr{H}$ since $\sum_n \frac{1}{n^2} < \infty$ by Proposition 9.3.1. Suppose $x \in \text{span}\{e_n\}$, say $x = \sum_{k=1}^{N} \alpha_k e_k$. We have

$$\left\langle \sum_{k=1}^{N} \alpha_k e_k, e_{N+1} \right\rangle = 0$$

, but

$$\langle x, e_{N+1} \rangle = \left\{ \sum_{n=1}^{\infty} \frac{1}{n} e_n, e_{N+1} \right\}$$

Define $T(y) = \langle y, e_{N+1} \rangle$. Let $S_M = \sum_{n=1}^{M} \frac{1}{n} e^n$, then $T(S_M) \xrightarrow{M \to \infty} T(x)$. Also, note that for $M \geq N+1$

$$T(S_M) = \sum_{n=1}^{M} \frac{1}{n} \langle e_n, e_{N+1} \rangle = \frac{1}{N+1}$$

Thus, let $M \to \infty$, we have

$$\left\{ \sum_{n=1}^{\infty} \frac{1}{n} e_n, e_{N+1} \right\} = \frac{1}{N+1}$$

this yields a contradiction.
Now consider any infinite orthonormal set Λ and select a countable subset $\{e_n\}_n$. Form x as above and suppose $x \in \text{span}(\Lambda)$, say $x = \sum_{j=1}^{M} \beta_j f_j, f_j \in \Lambda$. Choose $e_N \in \{e_n\}_n$ such that $e_N \{f_1, \cdots, f_M\}$. We shall obtain a similar contradiction. Thus, Λ cannot span $\mathscr{H}$. Since Λ is an arbitrary orthonormal set, we cann never span an infinite Hilbert space by an orthonormal set. ■

R Note that we have constructed an element $x \in \mathscr{H}$ that is a **countable** linear combination of the orthonormal vectors, but not a finite linear combinations.

Definition 9.4.1 — Complete Orthonormal Set. An orthonormal set $S \subseteq \mathscr{H}$ is called complete if whenever $\langle x, s\rangle = 0$ for every $s \in S$, then $x = 0$. By a basis for the Hilbert space $\mathscr{H}$ we mean a complete orthonormal set.

R What this is saying is like a maximal orthonormal set that we cannot append more elements to the set while maintaining orthonormality.

■ **Example 9.8** Take S to be the standard coordinate vectors $\{e_n\}_n$ in l^2. This is a basis since if $x = (x_n)_n \in l^2$ and $\langle x, e_n\rangle = 0$ for all n, then $x_n = 0$ for all n and that means $x = 0$. ■

■ **Example 9.9** If $\mathscr{H}$ is a finite dimensional Hilbert space, then any complete orthonormal set is an algebraic basis. To see this, first note that any orthonormal set is linearly independent and hence is necessarily finite. Assume the complete orthonormal set is the set of vectors $\{e_j\}_{j=1}^N$. We just need to show it spans. Suppose not, then there is $x \notin \operatorname{span}\left(\{e_j\}_{j=1}^N\right)$. Let

$$y = x - \sum_{k=1}^{N} \langle x, e_k\rangle e_k$$

Then $y \neq 0$, but y is orthogonal to all $e_j, j = 1, \ldots, N$. Then, $\{e_j\}_{j=1}^N$ is not a complete orthonormal set. Contradiction. ■

R More generally, it is true that if S is a complete orthonormal set in $\mathscr{H}$, then $\mathscr{H}$ is the closure of the span of S, say $\mathscr{H} = \overline{\operatorname{span}(S)}$. The following theorem address the countable S case.

Theorem 22

1. If $\{e_k\}_{k=1}^\infty$ is a complete orthonormal set and $x \in \mathscr{H}$, then

$$x = \sum_{k=1}^{\infty} \langle x, e_k\rangle e_k$$

and therefore $\mathscr{H} = \overline{\operatorname{span}\{e_k\}_k}$.

Proof. Let $x \in \mathscr{H}$, by Bessel's inequality, we already have

$$\sum_{k=1}^{\infty} |\langle x, e_k\rangle|^2 \leq \|x\|^2 < \infty$$

Then, we know that $y = \sum_{k=1}^\infty \langle x, e_k\rangle e_k \in \mathscr{H}$. Then, $\langle y, e_k\rangle = \langle x, e_k\rangle, \forall k$. Thus, $(y - x) \perp e_k, \forall k$. As $\{e_k\}_k$ is complete, we have $y = x$. ■

2. If S is an orthonormal set and $\mathscr{H} = \overline{\operatorname{span}(S)}$, then S is complete.

Proof. Let $x \in \mathscr{H}$ be orthogonal to any vector in S. Then, $\langle x, y\rangle = 0, \forall y \in S$. As $x \in \overline{\operatorname{span}(S)}$, and $x = \lim_n x_n$ for $x_n \in \operatorname{span}(S)$. Thus, $0 = \langle x, x_n\rangle \to \langle x, x\rangle$. Thus, $x = 0$, which means S is complete. ■

Corollary 9.4.2 Let $\{e_n\}_n$ be orthonormal. Then $\{e_n\}$ is complete if and only if $\|x\|^2 = \sum |\langle x, e_k \rangle|^2$ for all $x \in \mathscr{H}$. (i.e. Bessel's inequality is an equality)

Proof. If $\{e_n\}$ is complete. By 1 of the Theorem 22, we have

$$x = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$$

and we apply Pythogorean theorem to get

$$\|x\|^2 = \sum_{k=1}^{\infty} \|\langle x, e_k \rangle\|^2$$

For the converse, we use contradiction. Suppose that $\{e_n\}$ is not complete while known that

$$\|x\|^2 = \sum_{k=1}^{\infty} \|\langle x, e_k \rangle\|^2$$

is true. Then, we can find an element $0 \neq x \in \mathscr{H}$ such that $\langle x, e_k \rangle = 0, \forall k$. But that means $\|x\|^2 = 0 \iff x = 0$. This yields a contradiction. ∎

Corollary 9.4.3 The orthonormal set $\{e_n\}_n$ is complete if and only if the linear map $T : \mathscr{H} \to l^2$ given by $T(x) = (\langle x, e_k \rangle)_{k=1}^{\infty}$ is injective.

Proof. Assume $\{e_n\}$ is complete and suppose $T(x) = T(y)$. Then $\langle x, e_k \rangle = \langle y, e_k \rangle$ for all k, so $(x - y) \perp e_k, \forall k$. Thus, $x = y$. The converse is similar. ∎

Exercise 9.3 Suppose $\{e_n\}_{n=1}^{\infty}$ is a complete orthonormal set in $\mathscr{H}$. Show that the map T of corollary 9.4.3 is an isometric isomorphism. (bijective, linear, and preseves the norm.

Proof.
1. **Injective:** shown in corollary 9.4.3
2. **Surjective:** Let $x = \{x_n\}_{n=1}^{\infty} \in l^2$ where $x_k \in \mathbb{C}$ and $\sum_{n=1}^{\infty} |x_k|^2 < \infty$. By Proposition 15 in 1HibertSpacesIntro proved in class, we have some $y \in \mathscr{H}$ such that $\langle y, e_n \rangle = x_n$ for all n and $\|y\| = \|(x_n)\|_{l^2}$. Thus,

$$T(y) = (\langle y, e_n \rangle)_{n=1}^{\infty} = \{x_n\}_{n=1}^{\infty}$$

   thus, T is surjective.
3. **Linear:** let $a, b \in \mathbb{C}$, then, for any $x, y \in \mathscr{H}$

$$T(ax + by) = (\langle ax + by, e_n \rangle)_{n=1}^{\infty}$$

   for each n, by the inner product property, we have

$$\langle ax + by, e_n \rangle = a \langle x, e_n \rangle + b \langle y, e_n \rangle$$

   thus,

$$(\langle ax + by, e_n \rangle)_{n=1}^{\infty} = (a \langle x, e_n \rangle + b \langle y, e_n \rangle)_{n=1}^{\infty} = aT(x) + bT(y)$$

4. **Preserve inner product:** let $x, y \in \mathscr{H}$. Since $\{e_n\}_{n=1}^{\infty}$ is a complete orthonormal set for $\mathscr{H}$, by Theorem 7 in 2HilbertSpacesBasis proved in class, we have

$$x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$$

and

$$y = \sum_{n=1}^{\infty} \langle y, e_n \rangle e_n$$

Then, we start with the inner product in l^2 (labeled by l^2 in subscript, otherwise, we mean the inner product on $\mathscr{H}$)

$$\begin{aligned} \langle T(x), T(y) \rangle_{l^2} &= \langle (\langle x, e_n \rangle)_{n=1}^{\infty}, (\langle y, e_n \rangle)_{n=1}^{\infty} \rangle_{l^2} \\ &= \sum_{n=1}^{\infty} \langle x, e_n \rangle \overline{\langle y, e_n \rangle} \end{aligned}$$

from the other direction, we have

$$\langle x, y \rangle = \left\langle \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n, \sum_{n=1}^{\infty} \langle y, e_n \rangle e_n \right\rangle$$

for any $N \in \mathbb{N}$, we have

$$\left\langle \sum_{n=1}^{N} \langle x, e_n \rangle e_n, \sum_{n=1}^{N} \langle y, e_n \rangle e_n \right\rangle = \sum_{n=1}^{N} \langle x, e_n \rangle \overline{\langle y, e_n \rangle} \langle e_n, e_n \rangle = \sum_{n=1}^{N} \langle x, e_n \rangle \overline{\langle y, e_n \rangle}$$

this is true since $\{e_n\}_{n=1}^{\infty}$ is orthonormal and inner product has that sesquilinearity. Taking $N \to \infty$, we have

$$\langle x, y \rangle = \sum_{n=1}^{\infty} \langle x, e_n \rangle \overline{\langle y, e_n \rangle} = \langle T(x), T(y) \rangle$$

as previously deduced. Thus, T preserves the inner product.

■

Let's look at some major theorems about Hilbert spaces.

Theorem 23 Every Hilbert space has a basis. If $\mathscr{H}$ is a separable Hilbert space, then any orthonormal set is countable or finite.

Proof. Let $\mathscr{S}$ be the collection of all orthonormal sets from $\mathscr{H}$. Define the partial order by inclusion on $\mathscr{S}$.
Let $\mathscr{C}$ be chain from $\mathscr{S}$. Take $X = \cup_{C \in \mathscr{C}} C$. Certainly this dominates everything in the certain. We want to show $X \in \mathscr{S}$. Consider $x \neq \in X$. Since $\mathscr{C}$ is a chain, there is some $C \in \mathscr{C}$ such that contains both x, y. Since C is an orthonormal set, we have $\|x\| = \|y\| = 1$ and $\langle x, y \rangle = 0$. Thus, X is an orthonormal set.
By Zorn's Lemma, $\mathscr{S}$ has a maximal element. Call this maximal element A. Being in $\mathscr{S}$, A is an orthonormal set. If it is not complete, then there is some $x \in \mathscr{H}, x \neq 0$, such that $\langle x, a \rangle = 0, \forall a \in A$. But then $A \cup \{x/\|x\|\}$ is still an orthornormal set and it strictly contains A, violating that A is maximal.

Hence A is a complete orthonormal set and therefore the Hilbert space has a basis.
Now, suppose $\mathscr{H}$ is a separable Hilbert space and that $\{e_\lambda\}_{\lambda\in\Lambda}$ is a uncountable orthonormal set. Choose a countable dense subset $\{f_n\}_{n=1}^{\infty}$ of $\mathscr{H}$.
For each $\lambda \in \Lambda$, there is some $n \in \mathbb{N}$ with $e_\lambda \in B(f_n, 1/2)$. Since there are only countably many such balls and uncountably many e_λ, some ball, $B(f_N, 1/2)$, contains uncountably many e_λ.
In particular, there is a pair $e_1, e_2 \in B(f_N, 1/2)$. By the Pythogorean theorem, we have

$$\|e_1 - e_2\|^2 = \|e_1\|^2 + \|e_2\|^2 = 2$$

But, then

$$\begin{aligned}\sqrt{2} = \|e_1 - e_2\| &\le \|e_1 - f_n\| + \|f_n - e_2\| \\ &\le \frac{1}{2} + \frac{1}{2} = 1\end{aligned}$$

This yields a contradiction and we are done. ■

Corollary 9.4.4 Every separable infinite dimensional Hilbert space is isometrically isomorphic to l^2.

Proof. Delayed. ■

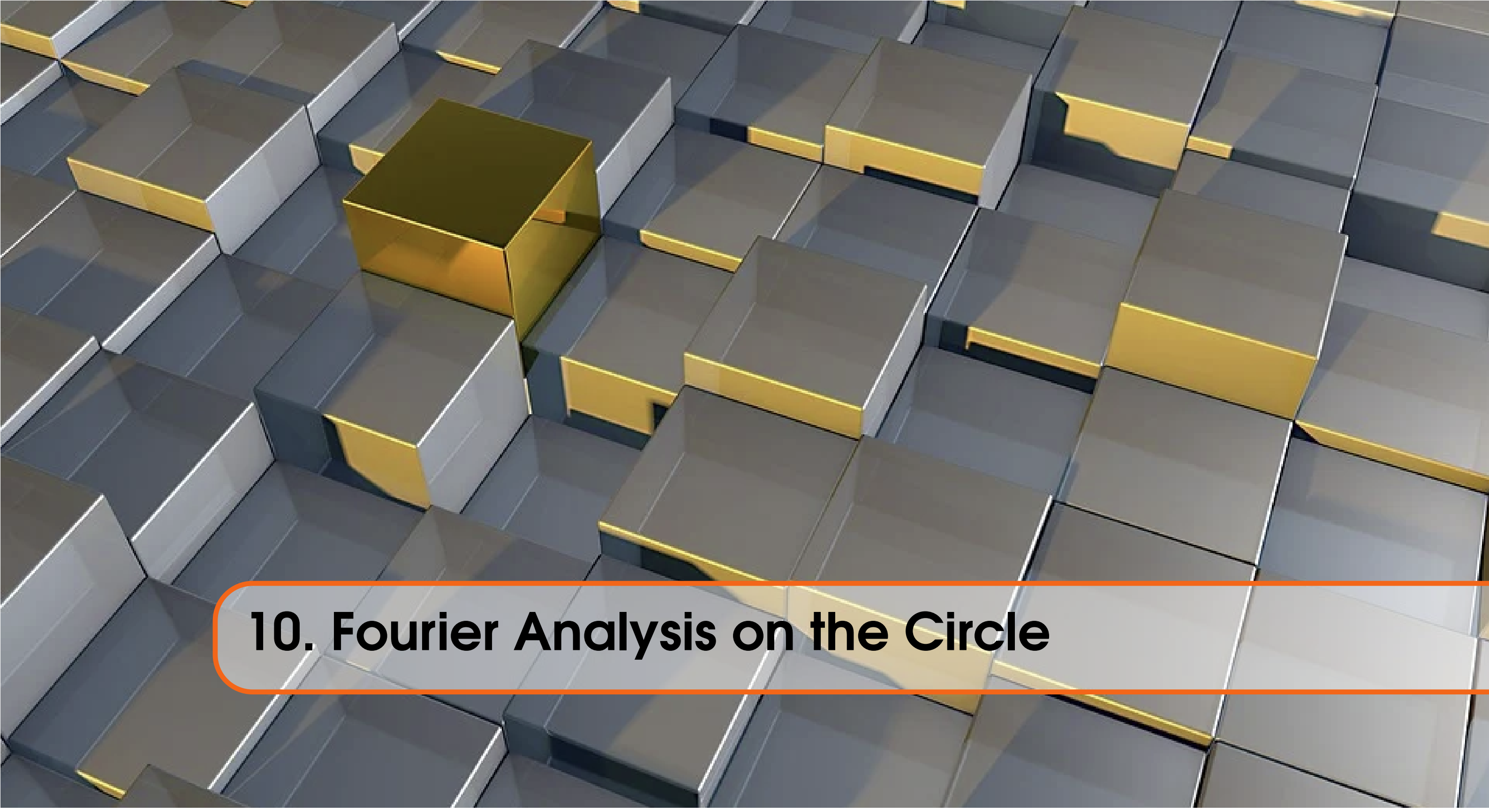
R In fact, the isometric isomorphism T can even be chosen to have the property that $\langle x, y\rangle = \langle T(x), T(y)\rangle$ for all $x, y \in \mathscr{H}$ where the first inner product is the inner product on $\mathscr{H}$ and the second inner product is the inner product on l^2. We say that a T is inner product preserving.

II Fourier Analysis

10 Fourier Analysis on the Circle 75
10.1 History
10.2 Basic Notation
10.3 Fourier Series and Fourier Coefficients
10.4 Parseval's Theorem
10.5 More About the Fourier Coefficients of L^1 Functions
10.6 Basic Properties of the Fourier Coefficients

11 The Dirichlet Kernel 83
11.1 Dirichlet Kernel
11.3 Intuitive Idea Why We Might Hope For $S_N f(x) \to f(x)$

12 Failure of Convergence of Fourier Series 87
12.1 Functional Analysis Ideas
12.2 Application - An L^1 Function with a Divergent Fourier Series
12.3 Application - A continuous function with a Divergent Fourier Series



10. Fourier Analysis on the Circle

10.1 History

Fourier analysis was named after Joseph Fourier, 1768-1830, a French mathematician who did much pioneering work. He was a French revolutionary who spent time in prison, as well as serving in Napoleon's government.
Fourier analysis is important today, both for theoretical and practical purposes. It has fundamental engineering applications in signal and image processing, including medical imaging, and in the study of partial differential equations.

10.2 Basic Notation

We will start our analysis on the circle group in $\mathbb{C}$ or the space $\mathbb{R}^n$. We denote the circle by $\mathbb{T}$,

$$\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$$

if we identify z with e^{it}, we can identify $\mathbb{T}$ as

$$\mathbb{T} = [0, 2\pi]$$

or we can shift it by 2π. Alternatively, by identifying $z = e^{2\pi it}$, we can view

$$\mathbb{T} = [0, 1]$$

or we can shift it by 1. We give $\mathbb{T}$ the topology it inherits from $\mathbb{C}$ and hence it is compact. It is also a metric space with $d_{\mathbb{T}}$.
Functions defined on $\mathbb{T} = [0, 2\pi]$ must satisfy $f(0) = f(2\pi)$ and hence have a unique $2\pi-$periodic extension to all of $\mathbb{R}$. By $C(\mathbb{T})$, we will mean all the continuous functions defined on $\mathbb{T}$.

Definition 10.2.1 We write m to denote the Lebesgue measure restricted to $[0, 2\pi)$ and normalized so $m(\mathbb{T}) = 1$, then

$$\int_{\mathbb{T}} f dm = \frac{1}{2\pi} \int_{[0,2\pi]} f(x)dx$$

For $1 \leq p < \infty$, we will put

$$\|f\|_{L^p(\mathbb{T})} = \|f\|_p = \left(\frac{1}{2\pi} \int_{[0,2\pi]} |f(x)|^p dx \right)^{1/p}$$

and let $\|f\|_\infty$ be the L^∞ norm on $\mathbb{T}$. Then, by our previous work, one can easily show that

$$C(\mathbb{T}) \subsetneq L^\infty(\mathbb{T}) \subsetneq L^p(\mathbb{T}) \subsetneq L^1(\mathbb{T}), 1 < p < \infty$$

and

$$\|f\|_1 \leq \|f\|_p \leq \|f\|_\infty$$

and an analog of the Hölder's Inequality,

$$\left| \frac{1}{2\pi} \int_{[0,2\pi]} fg \right| \leq \|f\|_p \|g\|_q$$

where (p,q) is a conjugate pair. In particular, when $p = q = 2$, we have that $L^2(\mathbb{T})$ is a Hilbert space with inner product

$$\langle f, g \rangle = \frac{1}{2\pi} \int_{[0,2\pi]} f\overline{g}$$

Exercise 10.1 Prove that $C(\mathbb{T})$ is dense in L^p for $p < \infty$. And prove that L^2 is dense in L^1.
Solution ■

10.3 Fourier Series and Fourier Coefficients

Proposition 10.3.1 The set of functions $\{\exp(inx) : n \in \mathbb{Z}\}$ is an orthonormal family of continuous functions on $\mathbb{T}$.

Proof.

$$\|e^{inx}\|_2^2 = \frac{1}{2\pi} \int_{[0,2\pi]} |e^{inx}|^2 dx = 1, \forall n \in \mathbb{Z}$$

and

$$\frac{1}{2\pi} \int_{[0,2\pi]} e^{inx}\overline{e^{imx}} dx = \frac{1}{2\pi} \int_{[0,2\pi]} e^{i(n-m)x} dx = \frac{1}{2\pi} \left[\frac{e^{i(n-m)x}}{i(n-m)} \bigg|_0^{2\pi} \right] = 0, n \neq m$$

■

R This is a countable family of orthonormal functions.

Definition 10.3.1 — Trigonometric Polynomial. A continuous function of the form

$$p(x) = \sum_{n=-M}^{N} a_n e^{inx}$$

The set of integers n where $a_n \neq 0$ are known as the **frequencies** of p. The **degree** of p is $\max |n|$ such that $a_n \neq 0$.

We write Trig$(\mathbb{T})$ for the set of all trigonometric polynomials. Thus,

$$\operatorname{Trig}(\mathbb{T}) := \operatorname{span}\{e^{inx} : n \in \mathbb{Z}\}$$

We have a really powerful theorem below.

Theorem 24 Trig$(\mathbb{T})$ is dense in $C(\mathbb{T})$ in the sup norm and $L^p(\mathbb{T})$ in the L^p norm for $p < \infty$.

Proof. Trig$(\mathbb{T})$ is dense in $C(\mathbb{T})$ is from an important result from PMATHH351, the Stone-Weiestrass Theorem. Note that Trig$(\mathbb{T})$ is a subalgebra in $C(\mathbb{T})$ that contains the constants (degree 0 trigonometric polynomials), separates points since if $s \neq t \mod 2\pi$, then $p(x) = e^{ix}$ has $p(s) \neq p(t)$, and is closed under conjugation. Thus, we have all the required conditions of the Stone-Weierstrass Theorem.
Since $C(\mathbb{T})$ is dense in $L^p(\mathbb{T})$ for $p < \infty$ and $\|\cdot\|_p \leq \|\cdot\|_\infty \leq$ sup norm, the Trig$(\mathbb{T})$ is also dense in $L^p(\mathbb{T})$. ■

Corollary 10.3.2 The span$\{e^{inx} : n \in \mathbb{Z}\}$ is dense in $L^2(\mathbb{T})$. Also, we know that $\{e^{inx} : n \in \mathbb{Z}\}$ is a complete orthonormal set and a basis for the Hilbert space $L^2(\mathbb{T})$.

R If the choice $\mathbb{T} = [0,1]$ is made, then $\{e^{i2\pi nx} : n \in \mathbb{Z}\}$ is a complete orthonormal set.

Definition 10.3.2 — Fourier Coefficient. The $n-$th Fourier coefficient of $f \in L^2(\mathbb{T})$ is

$$\hat{f}(n) = \langle f, e^{inx} \rangle = \frac{1}{2\pi} \int_{[0,2\pi]} f(x) e^{-inx} dx \in \mathbb{C}$$

the formal series

$$\sum_{n=-\infty}^{\infty} \hat{f}(n) e^{inx}$$

is known as the **Fourier series** of f.

10.4 Parseval's Theorem

Here are some of the important consequences of $\{e^{inx}\}_{n=-\infty}^{\infty}$ being a complete orthonormal set for $L^2(\mathbb{T})$.

1. **L^2 convergence of Fourier series to f:** when $f \in L^2(\mathbb{T})$, $\hat{f}(n) = \langle f, e^{inx} \rangle$ in our inner product notation. Thus, for such f, the Fourier series has the form

$$\sum_{n=-\infty}^{\infty} \langle f, e^{inx} \rangle e^{inx} \to f \in \mathscr{H} = \mathscr{L}^{\in}(\mathbb{T})$$

Put

$$S_N f(x) = \sum_{n=-N}^{N} \hat{f}(n) e^{inx}$$

This is a degree $\leq N$ trigonometric polynomial. As Trig$(\mathbb{T})$ is dense in $L^2(\mathbb{T})$, we know that for each $f \in L^2$, there is a sequence of trignometric polynomials $p_n \to f$ in L^2 norm. Thus,

$$f = \sum_{n=-\infty}^{\infty} \hat{f}(n) e^{inx}$$

the equality is in the L^2 sense.

2. **Uniqueness of Fourier Coefficients:** if $f \in L^2$ and $\hat{f}(n) = \langle f, e^{inx} \rangle = 0$ for all $n \in \mathbb{Z}$, then $f = 0$ a.e. This is immediate from the definition of a complete orthonormal set. Thus, if $f, g \in L^2$ and $\hat{f}(n) = \hat{g}(n)$ for all $n \in \mathbb{Z}$, then $\widehat{(f-g)}(n) = 0$ for all n and therefore $f = g$ a.e., so the Fourier coefficients uniquely determines the function.

Theorem 25 — Parseval's Theorem. For all $f \in L^2$,

$$\|f\|_2 = \left(\frac{1}{2\pi}\int_{[0,2\pi]|f|^2}\right)^{1/2} = \left(\sum_{n=-\infty}^{\infty} \left|\hat{f}(n)\right|^2\right)^{1/2} = \left\|\left(\hat{f}(n)\right)\right\|_{l^2(\mathbb{Z})}$$

since complete orthonormal set of $\left\{e^{inx}\right\}$ gives us

$$\|f\|_2 = \left(\sum \left|\langle f, e^{inx} \rangle\right|^2\right)$$

by Bessel's inequality's equality case.

R. Thus, the L^2 norm of f and the l^2 norm of the sequence of Fouier coefficients of f coincide. Another way of saying this is that the map

$$\underset{f \mapsto \left(\hat{f}(n)\right)_{n\in\mathbb{Z}}}{T : L^2(\mathbb{T}) \to l^2(\mathbb{Z})}$$

Theorem 26 — Upgraded Parseval's Theorem. The introduced map T is inner product preserving. Thus,

$$\langle f, g \rangle_{L^2} = \langle T(f), T(g) \rangle_{l^2}$$

or say

$$\frac{1}{2\pi}\int_{[0,2\pi]} f\overline{g} = \sum_{n=-\infty}^{\infty} \hat{f}(n)\overline{\hat{g}(n)}, \forall f, g \in L^2$$

taking $f = g$ gives the original Parseval's Theorem.

Exercise 10.2 Suppose $f \in L^2$ has Fourier series

$$1 + \sum_{n=1}^{\infty} \frac{(-1)^n}{n}\cos(nx) = 1 + \sum_{n=1}^{\infty} \frac{(-1)^n}{n}\left(\frac{e^{inx} + e^{-inx}}{2}\right)$$

1. Find $\hat{f}(0)$ and $\hat{f}(-2)$
2. Determine $\|f\|_2$

Solution:

1. $\hat{f}(0) = 1$ and $\hat{f}(-2) = \frac{1}{4}$, we cans see that pattern that

$$\hat{f}(n) = \begin{cases} 1 & n = 0 \\ \frac{(-1)^{|n|}}{2|n|} & |n| > 0 \end{cases}$$

2. By Parseval's Theorem, we have

$$\|f\|_2 = \left(\sum_{-\infty}^{\infty} \left|\hat{f}(n)\right|^2\right)^{1/2} = \left(1 + \frac{1}{2}\sum_{n=1}^{\infty}\frac{1}{n^2}\right)^{1/2} = \left(1 + \frac{1}{2}\frac{\pi^2}{6}\right)^{1/2}$$

■

Exercise 10.3 Let $a_n = |n|^{1/3}$ for $n \neq 0$. Is there a continuous function f such that $\hat{f}(n) = a_n$ for all $n \neq 0$? Explain
Solution: ■

10.5 More About the Fourier Coefficients of L^1 Functions

We have seen that $L^2(\mathbb{T})$ is a proper subset of $L^1(\mathbb{T})$. If $f \in L^1$, then $f(x)e^{-inx}$ is an integrable function, so we can still make the definition

$$\hat{f}(n) = \frac{1}{2\pi}\int_{[0,2\pi]} f(x)e^{-inx}dx, n \in \mathbb{Z}$$

we continue to call this the $n-$th Fourier coefficient of f. Note that $\hat{f}(n)$ is a complex number even if f is real-valued.
We lost quite some stuff since L^1 is not a Hilbert space. We want to explore whether those nice properties still exist.
We define the Fourier series of $f \in L^1$ as the formal sum

$$S(f) \sim \sum_{-\infty}^{\infty} \hat{f}(n)e^{inx}$$

Exercise 10.4 Write

$$\sum_{n=-\infty}^{\infty} a_n e^{0mx} = \frac{A_0}{2} + \sum_{n=1}^{\infty} A_n \cos(nx) + B_n \sin(nx)$$

where

$$A_n = a_n + a_{-n} = \frac{1}{\pi}\int_{[0,2\pi]} f(x)\cos(nx)dx$$

and

$$B_n = i(a_n - a_{-n}) = \frac{1}{\pi}\int_{[0,2\pi]} f(x)\sin(nx)dx$$

note that A_n, B_n are real-valued when f is real-valued. ■

10.6 Basic Properties of the Fourier Coefficients

Let $f, g \in L^1(\mathbb{T})$ and $n \in \mathbb{Z}$.

1. **Linearity:** $\widehat{f+g}(n) = \hat{f}(n) + \hat{g}(n)$, $\widehat{\alpha f}(n) = \alpha\hat{f}(n), \alpha \in \mathbb{C}$.
2. **Conjugation:** $\hat{\bar{f}}(n) = \overline{\hat{f}(n)}$ (check $\int \bar{g} = \overline{\int g}$).
3. Let $t \in \mathbb{T}$ and define $f_t(x) = f(x-t)$ where here the subtraction is understood mod 2π. Then,

$$\hat{f}_t(n) = e^{-int}\hat{f}(n)$$

Proof. Since $e^{-inx} = e^{-in(x-t)}e^{-int}$, we have

$$\begin{aligned}\hat{f}_t(n) &= \frac{1}{2\pi}\int_{[0,2\pi]} f_t(x)e^{-inx}dx = \frac{1}{2\pi}\int_{[0,2\pi]} f(x-t)e^{-inx}dx \\ &= \frac{e^{-int}}{2\pi}\int_{[0,2\pi]} f(x-t)e^{-in(x-t)}dx\end{aligned}$$

By the translation invariance of Lebesgue measure but mod 2π (not trivial). Hence,

$$\hat{f}_t(n) = e^{-int}\hat{f}(n)$$

■

4. $\left|\hat{f}(n)\right| \le \|f\|_1$. Thus, $\left(\hat{f}(n)\right)_{n\in\mathbb{Z}} \in l^\infty$ and $\left\|\left(\hat{f}(n)\right)\right\|_{l^\infty} \le \|f\|_1$. Here

$$\|(x_n)\|_{l^\infty} = \sup_n |x_n|$$

5. If $f_k \to f \in L^1$, then $\hat{f}_k(n) \to \hat{f}(n)$ for all $n \in \mathbb{Z}$ (even uniformly in n).

Proof.

$$\left|\hat{f}_k(n) - \hat{f}(n)\right| = \left|\widehat{f_k - f}(n)\right| \le \|f_k - f\|_1 \to 0$$

■

Exercise 10.5 With f_t, use the translation invariance property of the integral to prove that $\|f_t\|_p = \|f\|_p$ for all $t \in \mathbb{T}$ and $p \ge 1$. ■

Exercise 10.6 Prove $\lim_{t\to 0}\|f_t - f\|_p = 0$ for $1 \le p < \infty$. (Hint. do it for continuous function first) ■

Lemma 10.7 — Riemann-Lebesgue Lemma. Let $f \in L^1(\mathbb{T})$. Then, $\lim_{n\to\pm\infty}\hat{f}(n) = 0$.

Proof. Let $\varepsilon > 0$. Since the trigonometric polynomials are dense in all the L^p spaces for $p < \infty$, we can choose $p \in \mathrm{Trig}(\mathbb{T})$ such that $\|f - p\|_1 < \varepsilon$. Let $N = \deg p$. Then, $\hat{p}(n) = 0$ for all $|n| > N$. Consequently, for all $n > |N|$, by linearity and (4) we have

$$\left|\hat{f}(n)\right| \le \left|\hat{f}(n) - \hat{p}(n)\right| + |\hat{p}(n)| \le \|f - p\|_1 < \varepsilon$$

■

R Thus the sequence of Fourier coefficients $\left(\hat{f}(n)\right)_n$ actually belongs to the subspace of l^∞ known as $c_0 = \{(x_n)_n : |x_n| \to 0\}$.

Exercise 10.7 Verify that c_0 is a subspace of l^∞ and that it is closed. ■

R It is worth observing that this result is very easy for $f \in L^2$. The sequence $\left(\hat{f}(n)\right)_n$ satisfies $\sum_{n=-\infty}^{\infty}\left|\hat{f}(n)\right|^2 = \|f\|_2^2 < \infty$ by Parseval's Theorem and hence we certainly have $\hat{f}(n) \to 0$. Alternative proof of the Riemann-Lebesgue Lemma can be given by using the densenss of L^2 in L^1.

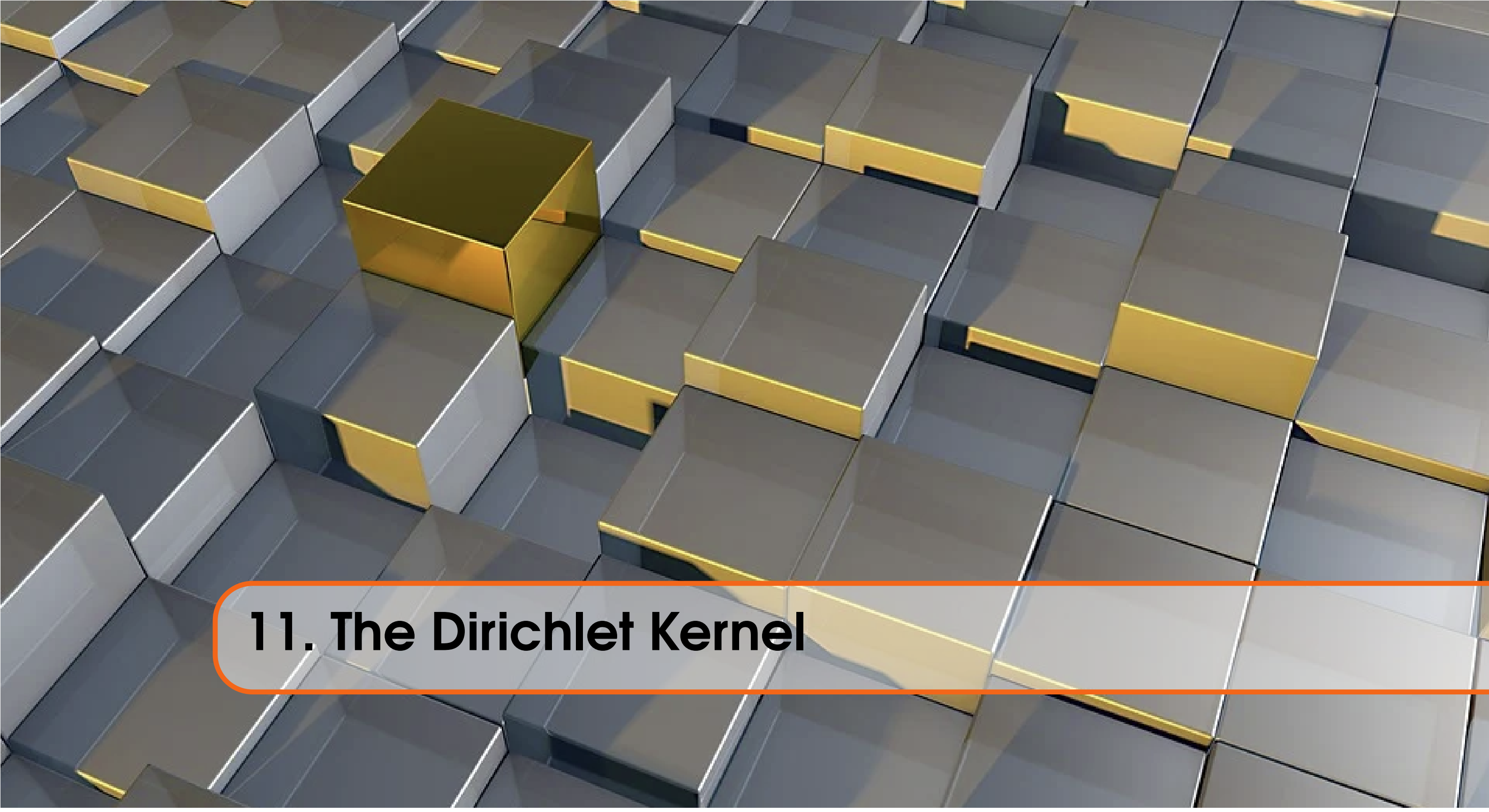
Corollary 10.7.1 If $f \in L^1(\mathbb{T})$, then

$$\lim_{n\to\infty}\int_{[0,2\pi]} f(x)\sin(nx)dx = 0 = \lim_{n\to\infty}\int_{[0,2\pi]} f(x)\cos(nx)dx$$

Exercise 10.8 If $f \in L^1(\mathbb{T})$, then for all β real,

$$\lim_{n\to\infty} \int_{[0,2\pi]} f(x)\sin(nx+\beta)\,dx = 0$$

■



11. The Dirichlet Kernel

11.1 Dirichlet Kernel

How can we recover a function through its Fourier series?

Definition 11.1.1 — N-th Dirichlet Kernel.

$$D_N(t) = \sum_{n=-N}^{N} e^{int} = 1 + 2\sum_{n=1}^{N} \cos(nt) \in \mathrm{Trig}(\mathbb{T})$$

R This is an even, real-valued, degree N trigonometric polynomial. Here are two other obvious facts

1. $D_N(0) = 2N + \sup_t |D_N(t)|$
2. $\widehat{D_N}(n) = 1$ if $|n| \leq N$ and 0 else

Lemma 11.2 The following formula holds for all $t \in \mathbb{T}$, where we understand it in the limiting sense when $t = 0$:

$$D_N(t) = \frac{\sin[(N+1/2)t]}{\sin(t/2)}$$

Proof.

$$\begin{aligned}
(\sin(t/2))D_N(t) &= \left(\frac{e^{it/2} - e^{-it/2}}{2i}\right) \sum_{n=-N}^{N} e^{int} \\
&= \frac{e^{i(N+1/2)t} - e^{-i(N+1/2)t}}{2i} = \sin[(N+1/2)t]
\end{aligned}$$

■

R A consequence of this identity is that if $t \in (\delta, 2\pi - \delta)$, then $|D_N(t)| \leq \frac{1}{|\sin \delta/2|}$ is a bound which is independent of the choice of N.

If $f \in L^1$, then

$$\begin{aligned}
S_N f(x) &= \sum_{n=-N}^{N} \hat{f}(n) e^{inx} = \sum_{n=-N}^{N} \left(\frac{1}{2\pi} \int_{[2,\pi]} e^{-int} dt \right) e^{inx} \\
&= \sum_{n=-N}^{N} \left(\frac{1}{2\pi} \int_{[2,\pi]} f(t) e^{in(x-t)} dt \right) = \frac{1}{2\pi} \int_{[0,2\pi]} f(t) \sum_{n=-N}^{N} e^{in(x-t)} dt \\
&= \frac{1}{2\pi} \int_{[0,2\pi]} f(t) D_N(x-t) dt = \frac{1}{2\pi} \int_{[0,2\pi]} f(t) D_N(t-x) dt && \text{even-ness} \\
&= \frac{1}{2\pi} \int_{[0,2\pi]} f(t+x) D_N(t) dt && \text{trans. invar.}
\end{aligned}$$

11.3 Intuitive Idea Why We Might Hope For $S_N f(x) \to f(x)$

Recall that $f_t(x) = f(x-t)$ and

$$\begin{aligned}
\|f_t\|_p &= \|f\|_p, \forall t \in \mathbb{T}, p \geq 1 \\
\hat{f}_t(n) &= e^{-int} \hat{f}(n), \forall n
\end{aligned}$$

then, from previous section, we have

$$\begin{aligned}
S_N f(x) &= \frac{1}{2\pi} \int_{[0,2\pi]} f(t+x) D_N(t) dt \\
&= \frac{1}{2\pi} \int_{[0,2\pi]} f_{-x}(t) \left(\frac{\sin[(N+1/2)t]}{\sin(t/2)} \right) dt \\
&= \frac{1}{2\pi} \int_{[0,2\pi]} f_{-x}(t) \left(\sin(Nt) \frac{\cos(t/2)}{\sin(t/2)} + \cos(Nt) \right) dt
\end{aligned}$$

Assume $f \in L^1$, then so is f_{-x}. Then, by the Riemann-Lebesgue Lemma,

$$\frac{1}{2\pi} \int_{[0,2\pi]} f_{-x}(t) \cos(Nt) dt \to 0, \text{ as } N \to \infty$$

We then focus on

$$\begin{aligned}
\frac{1}{2\pi} \int_{[0,2\pi]} f_{-x}(t) \left(\sin(Nt) \frac{\cos(t/2)}{\sin(t/2)} \right) dt &= \frac{1}{2\pi} \int_{[0,\delta] \cup [2\pi-\delta,2\pi]} f_{-x}(t) \left(\sin(Nt) \frac{\cos(t/2)}{\sin(t/2)} \right) dt \\
&\quad + \frac{1}{2\pi} \int_{[\delta,2\pi-\delta]} f_{-x}(t) \left(\sin(Nt) \frac{\cos(t/2)}{\sin(t/2)} \right) dt \\
&= I + J
\end{aligned}$$

we will select δ a bit later.

1. J:

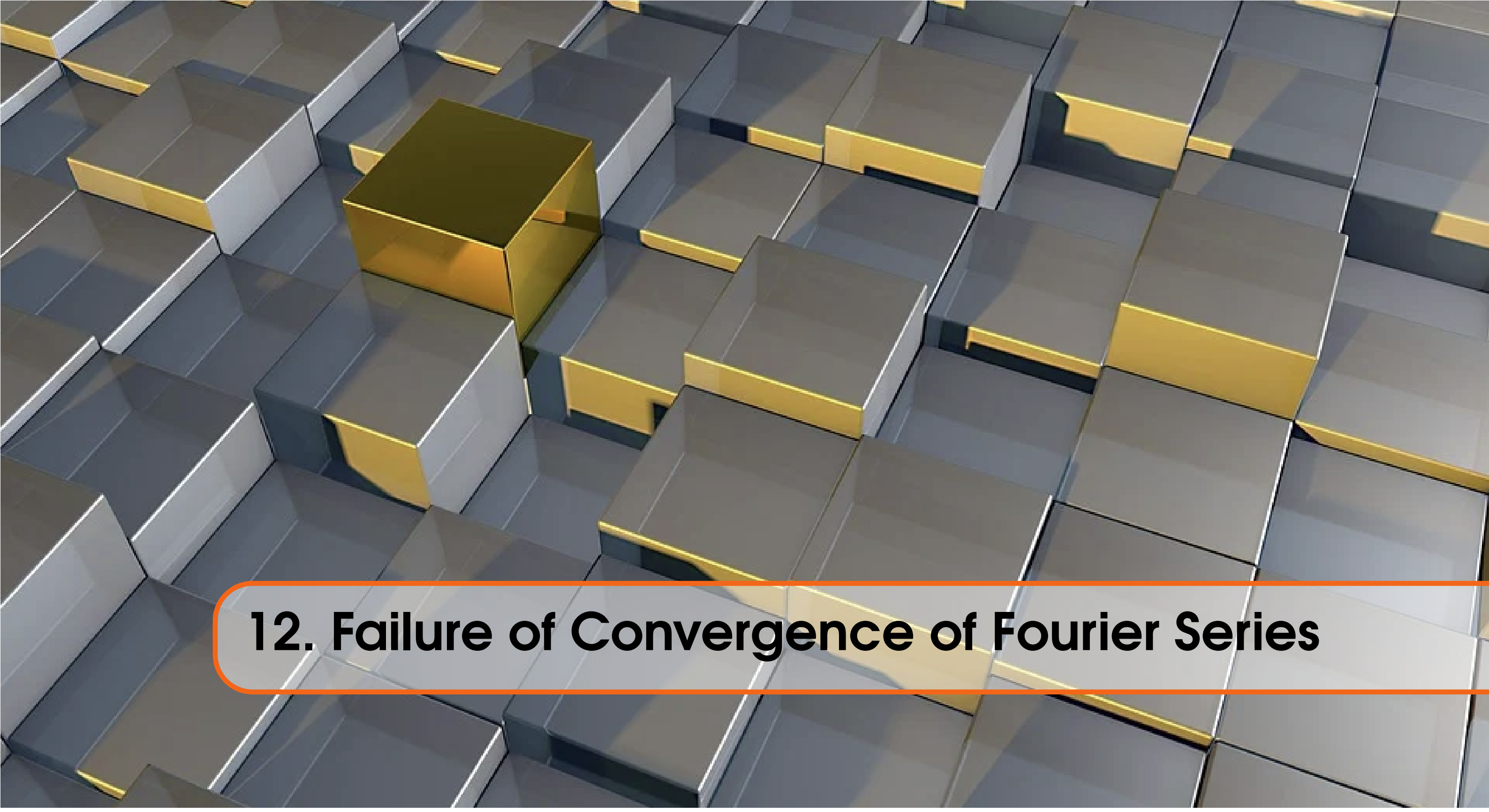
$$J = \frac{1}{2\pi} \int \underbrace{\left(f_{-x}(t) \chi_{[\delta,2\pi-\delta]} \frac{\cos(t/2)}{\sin(t/2)} \right)}_{g(x)} \sin(Nt) dt$$

   since $g \in L^1$, by Riemann-Lebesgue Lemma, we have

$$J \xrightarrow{N \to \infty} 0$$

2. I:
$$I = \frac{1}{2\pi} \int_{[-\delta,\delta]} f(x+t) \left(\sin(Nt) \frac{\cos(t/2)}{\sin(t/2)} \right) dt$$
where now we understand $\mathbb{T} = [-\pi, \pi]$. Then, for small enough $\delta > 0$, $f(x+t) \approx f(x)$ for $|t| < \delta$. Then,
$$\begin{aligned} I &\approx \frac{f(x)}{2\pi} \int_{[-\delta,\delta]} \sin(Nt) \frac{\cos(t/2)}{\sin(t/2)} dt \approx \frac{f(x)}{2\pi} \int_{\mathbb{T}} D_N(t) dt \\ &= f(x) \widehat{D_N}(0) = f(x) \end{aligned}$$

R In fact, it is NOT true that the Fourier series of a continuous function converges pointwise. (Altough it is most of the case...). We will give two approaches to show there exists some L^1 function with Fourier Series diverges in L^1.



12. Failure of Convergence of Fourier Series

12.1 Functional Analysis Ideas

Powerful results from functional analysis can give an existensial proof.

Definition 12.1.1 — Linear. Let X,Y be Banach spaces. A map $F : X \to Y$ is called linear if $F(\alpha x + y) = \alpha F(x) + F(y)$ for all $x, y \in X$ and scalars α. Given a linear map F we define

$$\|F\|_{op} := \sup_{\|x\|_X \leq 1} \|F(x)\|_Y$$

we call F bounded if $\|F\|_{op} < \infty$. We call $\{x \in X : \|x\|_X \leq 1\}$ is the closed unit ball of X.

R We have seen in one of the assignment questions that a closed unit ball of an infinite dimensional Banach space is never compact.

R For any linear map F, $F(0) = 0$. Also, by linearity,

$$\|F\|_{op} := \sup_{x \neq 0} \frac{\|F(x)\|_Y}{\|x\|_X}$$

Thus, $\|F\|_{op}\|x\|_X \geq \|F(x)\|_Y, \forall x \in X$.

Proposition 12.1.1 A linear map is bounded if and only if it is continuous.

Proof. Suppose F is bounded and let $x_n \to x$ in $\|\cdot\|_X$. Then

$$\|F(x_n) - F(x)\|_Y = \|F(x_n - x)\|_Y \leq \|F\|_{op}\|x_n - x\|_X \to 0$$

so F is continuous.
Conversely, suppose F is continuous at 0. Then for every $\varepsilon > 0$, there is some $\delta > 0$ such that $\|x\|_X \leq \delta$, then $\|F(x) - F(0)\|_Y < \varepsilon$. Apply this with $\varepsilon = 1$ and the corresponding δ. It $\|x\|_X \leq 1$,

then $\|\delta x\|_X \leq \delta$, so $\delta\|F(x)\|_Y = \|F(\delta x)\|_Y < 1$. Hence, $\|F(x)\|_Y \leq \frac{1}{\delta}$ whenever $\|x\|_Y \leq 1$ and that implies $\|F\|_{op} \leq \frac{1}{\delta}$ and bounded. ■

■ **Example 12.1** Fix an integer n and consider $F : L^(\mathbb{T}) \to \mathbb{C}$ given by $F(f) = \hat{f}(n)$. This is a linear map since $\widehat{\alpha f + g}(n) = \alpha\hat{f}(n) + \hat{g}(n)$. And since $\left|\hat{f}(n)\right| \leq \|f\|_1$, we have $\|F\|_{op} \leq 1$. Taking $f(x) = e^{inx}$, we have $\|f\|_1 = 1$ and $F(f) = 1$, hence $\|F\| = 1$, hence $\|F\| = 1$ (we omit the subscript op if this is clear) ■

■ **Example 12.2** Consider $F : L^2 \to l^2$ given by $F(f) = \left(\hat{f}(n)\right)_n$. This is an isometry: $\|F(f)\|_{l^2} = \|f\|_{L^2}$, hence $\|F\| = 1$. (Isometries are always bounded linear maps between Banach spaces) ■

■ **Example 12.3 — IMPORTANT.** Take $S_N : L^1 \to L^2$ defined by

$$S_N(f)(x) = \sum_{n=-N}^{N} \hat{f}(n)e^{inx} = \frac{1}{2\pi}\int_{[0,2\pi]} f(t+x)D_N(t)dt$$

the linearity is clear. We check that S_N is bounded.

$$\begin{aligned}
\|S_N f\|_1 &= \frac{1}{2\pi}\int_{\mathbb{T}}\left|\frac{1}{2\pi}\int_{\mathbb{T}} f(t+x)D_N(t)dt\right|dx \\
&\leq \frac{1}{2\pi}\int_{\mathbb{T}}\frac{1}{2\pi}\int_{\mathbb{T}} |f(t+x)D_N(t)dt|dx \\
&= \frac{1}{2\pi}\int_{\mathbb{T}} |D_N(t)|\left(\frac{1}{2\pi}\int_{\mathbb{T}} |f(t+x)|dx\right)dt \quad \text{Fubini}
\end{aligned}$$

since $\|f_{-t}\|_1 = \|f\|_1$, we have

$$\begin{aligned}
\|S_N f\|_1 &\leq \frac{1}{2\pi}\int_{\mathbb{T}} |D_N(t)|\left(\frac{1}{2\pi}\int_{\mathbb{T}} |f_{-t}(x)|dx\right)dt \\
&\leq\leq \frac{1}{2\pi}\int_{\mathbb{T}} |D_N(t)|\|f\|_1 dt = \|D_N\|_1\|f\|_1 \leq \|D_N\|_\infty\|f\|_1 \\
&= (2N+1)\|f\|_1
\end{aligned}$$

Hence, for $\|f\|_1 \leq 1$, we have $\|S_N\|_{op} \leq \|D_N\|_1 < \infty$. Thus, S_N is a bounded linear map for each N.

Claim: $\|S_N\|_{op} = \|D_N\|_1$

Proof. This will be useful for our divergence result later. Consider the function

$$f_k = \frac{\chi_{(-1/k,1/k)}}{m(-1/k,1/k)}$$

we are reminded that $m(-1/k,1/k) = 2/2k\pi = 1/k\pi$ and then $\|f_k\|_1 = 1$. We have

$$\begin{aligned}
\|S_N(f_k) - D_N\|_1 &= \frac{1}{2\pi}\int_{\mathbb{T}} |S_N(f_k)(x) - D_N(x)|dx \\
&= \frac{1}{2\pi}\int_{\mathbb{T}}\left|\frac{1}{2\pi}\int_{\mathbb{T}} f_k(t)D_N(t-x) - D_N(x)\right|dx
\end{aligned}$$

we note that

$$\frac{1}{2\pi}\int_{\mathbb{T}} f_k = 1$$

(iNtErEStinG trick)

$$\begin{aligned}\|S_N(f_k)-D_N\|_1 &= \frac{1}{2\pi}\int_{\mathbb{T}}\left|\frac{1}{2\pi}\int_{\mathbb{T}} f_k(t)D_N(t-x)-\left(\frac{1}{2\pi}\int_{\mathbb{T}} f_k(t)dt\right)D_N(x)\right|dx \\ &= \frac{1}{2\pi}\int_{\mathbb{T}}\left|\frac{1}{2\pi}\int_{\mathbb{T}} f_k(t)D_N(t-x)-\frac{1}{2\pi}\int_{\mathbb{T}} D_N(x)f_k(t)dt\right|dx \\ &= \frac{1}{2\pi}\int_{\mathbb{T}}\left|\frac{1}{2\pi}\int_{\mathbb{T}} f_k(t)(D_N(t-x)-D_N(x))dt\right|dx\end{aligned}$$

Since D_N is even and f_k is non-zero on $(-1/k,1/k)$, then,

$$\|S_N(f_k)-D_N\|_1 \le \frac{1}{2\pi}\int_{\mathbb{T}}\frac{1}{2\pi}\int_{-1/k}^{1/k}|f_k(t)(D_N(t-x)-D_N(x))|dtdx$$

suppose that we have $\varepsilon>0$. Using the uniform continuity of D_N, we can choose k such that $|D_N(x-t)-D_N|<\varepsilon$ for $|t|<\frac{1}{k}$ and for all x. Then,

$$\|S_N(f_k)-D_N\|_1 \le \frac{1}{2\pi}\int_{\mathbb{T}}\frac{1}{2\pi}\int_{-1/k}^{1/k}|f_k(t)\varepsilon|dtdx \le \frac{\varepsilon}{2\pi}\int_{\mathbb{T}}\|f_k\|_1 dx<\varepsilon$$

Thus, $S_N(f_k)\to D_N$ in L^1 as $k\to\infty$. Then,

$$\|S_N\|_{op} \ge \frac{\|S_N(f_k)\|_1}{\|f_k\|_1} \xrightarrow{k\to\infty} \|D_N\|_1$$

Thus, $\|S_N\|_{op}=\|D_N\|_1$. ■

■

Now, we introduce the key functional analysis result.

Theorem 27 — Uniform Boundedness Principle. Suppose X,Y are Banach spaces. Let $\mathscr{F}$ be a family of bounded linear maps from X to Y. If for every $x\in X$, $\sup_{F\in\mathscr{F}}\|F(x)\|_Y<\infty$, then $\sup_{F\in\mathscr{F}}\|F\|_{op}<\infty$. Consequently, if $\sup_{F\in\mathscr{F}}\|F\|_{op}=\infty$, then there is some $x\in X$ where $\sup_{F\in\mathscr{F}}\|F(x)\|_Y=\infty$.

12.2 Application - An L^1 Function with a Divergent Fourier Series

Theorem 28 There exists $f\in L^1$ such that $\sup_N\left\|\sum_{n=-N}^N \hat{f}(n)e^{inx}\right\|_1=\infty$. Thus, f has a divergence Fourier series.

Proof. Take $X=Y=L^1(\mathbb{T})$ and consider the family $\mathscr{F}=\left\{S_N:L^1\to L^2, N\in\mathbb{N}\right\}$ as in Example 12.3. We have seen that $\|S_N\|_{op}=\|D_N\|_1$.
We will show later that $\|D_N\|_1\ge C\log N$ for some $C>0$. Then, $\sup_{S_N\in\mathscr{F}}\|S_N\|_{op}=\infty$ and hence by the Uniform boundedness principle, there exists $f\in L^1$ such that $\sup_N\|S_N(f)\|_1=\infty$. But that means $\sup_N\left\|\sum_{n=-N}^N\hat{f}(n)e^{inx}\right\|_1=\infty$. ■

Proposition 12.2.1 There is a constant $C>0$ such that $\|D_N\|_1\ge C\log N$ for all N.

Proof. Being a trigonometric polynomial, each D_N is Riemann integrable. Hence all the following Lebesgue integrals can be understood as Riemann integrals. Then,

$$\|D_N\|_1=\frac{1}{2\pi}\int_0^{2\pi}\left|\frac{\sin[(N+1/2)t]}{\sin(t/2)}\right|dt$$

let $y = t/2$, we have

$$\|D_N\|_1 = \frac{1}{\pi}\int_0^{\pi}\left|\frac{\sin[(2N+1)y]}{\sin(y)}\right|dy \geq \frac{1}{\pi}\int_0^{\pi/2}\left|\frac{\sin[(2N+1)y]}{\sin(y)}\right|dy$$

Note that

$$1 \leq \left|\frac{y}{\sin y}\right| \leq \frac{\pi}{2}, y \in (0, \pi/2)$$

then,

$$\|D_N\|_1 \geq \frac{1}{\pi}\int_0^{\pi/2}\left|\frac{\sin[(2N+1)y]}{y}\right|dy$$

let $x = (2N+1)y$ and partition the integrand:

$$\|D_N\|_1 \geq \frac{1}{\pi}\int_{)}^{(2N+1)\pi/2}\left|\frac{\sin x}{x}\right|dx = \frac{1}{\pi}\sum_{j=0}^{2N}\int_{j\pi/2}^{(j+1)\pi/2}\left|\frac{\sin x}{x}\right|dx$$
$$= \frac{1}{\pi}\sum_{j=0}^{2N}\frac{1}{(j+1)\pi/2}\int_{j\pi/2}^{(j+1)\pi/2}|\sin x|dx$$

note that $\sin x$ is of constant sign over the interval $[j\pi/2, (j+1)\pi/2]$ and there is constant area under the curve as shown in the graph below.

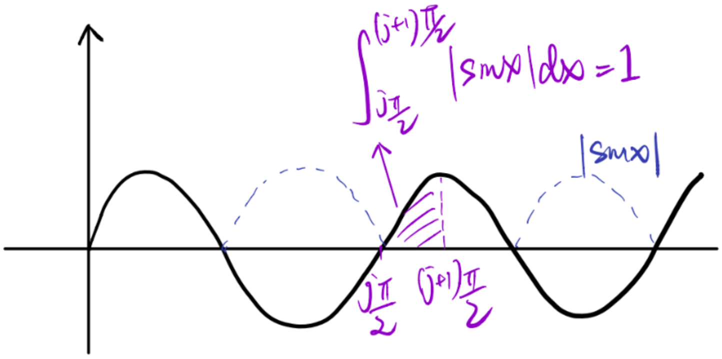


Figure 12.2.1: $|\sin(x)|$

Thus,

$$\|D_N\|_1 \geq \frac{1}{\pi}\sum_{j=0}^{2N}\frac{1}{(j+1)\pi/2} = \frac{2}{\pi^2}\sum_{j=1}^{2N+1}\frac{1}{j}$$
$$\geq \frac{2}{\pi^2}\int_1^{2N+1}\frac{1}{x}dx = \frac{2}{\pi^2}\log(2N+1) \geq C\log N$$

■

12.3 Application - A continuous function with a Divergent Fourier Series

Theorem 29 There is a continuous function f with $(S_n(f)(0))_n$ a divergent sequence.

Proof. HOLY SHIT THE PROOF IS LONG. (jks, will include later) ■