MATH 355: Honours Analysis 4

Pr. John Toth

Course notes by
Léo Raymond-Belzile
Leo.Raymond-Belzile@mail.mcgill.ca

The current version is that of April 23, 2013
Winter 2013, McGill University

Please signal the author by email if you find any typo.

These notes have not been revised and should be read carefully.

Contents

Section 1
Measure theory in $\mathbb{R}^n$

The basic question: given a rather "rough" set $E \subset \mathbb{R}^n$, how does one assign a "volume" to E, denoted by $|E| \equiv \mathrm{vol}(E)$. Some basic notions and definitions: For $\boldsymbol{x} \in \mathbb{R}^n$, denote $\boldsymbol{x} = (x_1, \ldots, x_n)$, (Euclidian coordinates), $x_j \in \mathbb{R}$ and $|\boldsymbol{x}| = \sqrt{x_1^2 + \cdots + x_n^2}$.

Set theory notation

Given $E \subset \mathbb{R}^n$, the complement of E is $E^{\complement} = \mathbb{R}^n - E$. The set $E - F = \{x \in F^{\complement} | x \notin F\}$.

Basic point-set topology

Given $r > 0$, we let $B_r(\boldsymbol{x}) = \{\boldsymbol{y} \in \mathbb{R}^n : |\boldsymbol{y} - \boldsymbol{x}| < r\}$ be the ball of radius $r > 0$ centered at $\boldsymbol{x} \in \mathbb{R}^n$

Definition 1.1

1. $E \subset \mathbb{R}^n$ is **open** if for every $\boldsymbol{x} \in E$, there exists $B_r(\boldsymbol{x}) \subset E$.
2. $E \subset \mathbb{R}^n$ is **closed** if $E^{\complement}$ is open, $E^{\complement} := \mathbb{R}^n - E$.
3. $E \subset \mathbb{R}^n$ is **bounded** provided there exist $B_R(\boldsymbol{x})$ with $R < \infty$ and $E \subset B_R(\boldsymbol{x})$.
4. $E \subset \mathbb{R}^n$ is **compact** if it is closed and bounded. By Heine-Borel, this is equivalent to the following property:

   Proposition 1.2 (Heine-Borel)
   Given $E \subset \bigcup_\alpha O_\alpha$, O_α open, there exist a finite subcover $O_{\alpha_1}, \ldots, O_{\alpha_N}$ open with $E \subset \bigcup_{i=1}^N O_{\alpha_i}$

5. $\boldsymbol{x} \in \mathbb{R}^n$ is a **limit point** of E if $B_r(\boldsymbol{x}) \cap E^{\complement} \neq \emptyset$ for every ball $B_r(\boldsymbol{x})$. Denote $lp(E) := \bigcup_{\boldsymbol{x}} \{\boldsymbol{x} \text{ is a limit point of } E\}$.
6. $\boldsymbol{x} \in E$ is an **interior point** if there exists $B_r(\boldsymbol{x}) \subset E$ for some $r > 0$

$$\mathbf{int}(E) = \bigcup_{\boldsymbol{x}} \{\boldsymbol{x} \text{ is an interior point of } E\}$$

7. The **closure** of E is given by $\overline{E} = \mathbf{int}(E) \cup lp(E)$
8. ∂E, the boundary of E, is defined as $\partial E := \overline{E} - \mathbf{int}(E)$.

1.1 Rectangles and cubes in $\mathbb{R}^n$

Denote $R = [a_1, b_1] \times \cdots \times [a_n, b_n] = \prod_{i=1}^n [a_i, b_i]$ where $a_i \leq b_i$ (the closed rectangle). The volume of R is given by

$$|R| = \prod_{j=1}^n |b_j - a_j| = \prod_{j=1}^n (b_j - a_j).$$

The corresponding open rectangle is given by $R = (a_1, b_1) \times \cdots \times (a_n, b_n)$ and again the volume $|R| = \prod_{j=1}^n (b_j - a_j)$.

Definition 1.3

A union of rectangles is almost disjoint if the interiors are disjoint

Lemma 1.4

Let $R = \bigcup_{k=1}^N R_k$ be an almost disjoint union of rectangles. Then $|R| = \sum_{k=1}^N |R_k|$.

By extending sides indefinitely as in picture, $R = \bigcup_{j=1}^M \tilde{R}_j$, $R_k = \bigcup_{j \in I_k} \tilde{R}_j$ where the $\tilde{R}_j$ are almost disjoint. Moreover, $|R| = \sum_{j=1}^M |\tilde{R}_j| = \sum_{k=1}^N \sum_{j \in I_k} |\tilde{R}_j|$ since the boundary faces of $\tilde{R}_j$ form a partition.

Lemma 1.5

Let $R_1, \ldots, R_N$ be rectangles and $R \subset \bigcup_{j=1}^N R_j$ is another rectangle. Then

$$|R| \leq \sum_{j=1}^N |R_j|$$

Theorem 1.6

Let $O \subset \mathbb{R}$ be open. Then $O = \bigcup_{j=1}^\infty I_j$ where I_j are disjoint open interval.

Proof. Given $x \in O$, we let I_x be the maximum open interval in O containing x. We have $I_x = (a_x, b_x)$ where

$$a_x = \sup\{a < x, (a, x) \in O\}$$
$$b_x = \inf\{b > x, (x, b) \in O\}$$

Claim

$O = \bigcup_{x \in O} I_x$ is clear.

Claim

I_x are disjoint.

Proof. Suppose not, $I_y \cap I_x \neq \emptyset$. Since $I_x \cup I_y \subset O$ and $x \in I_x \cup I_y$, then $I_x \cup I_y \subset I_x$ since I_x is maximal. Similarly, $I_x \cup I_y \subset I_y$ since I_y is maximal, which imply $I_x = I_y$.

So the union in Claim 1 is disjoint. To see that this union is countable, we just note that $I_x \cap \mathbb{Q} \neq \emptyset$. ■

■

Remark

When $n = 1$ and O is open, $|O| = \sum_{j=1}^{\infty} I_j$ where $\{I_j\}_{j=1}^{\infty}$ is the maximal covering in $\mathbb{R}^n$. Two issues not resolved: more general measurable sets in $\mathbb{R}$ and higher dimensions.

Last time, we proved that any open $O \subset \mathbb{R}$ can be written as a countable union of disjoint intervals $O = \bigcup_{j=1} I_j$; therefore we can define $|O| = \sum_{j=1}^{\infty} |I_j|$, where $|\cdot|$ is the "volume" or measure. Unfortunately, the situation in higher dimension is not so easy.

Theorem 1.7

Let $O \subset \mathbb{R}^n$ for $n \geq 1$ be open. Then there exists $\{Q_j\}_{j=1}^{\infty}$ almost disjoint cubes with the property that $O = \bigcup_{j=1}^{\infty} Q_j$.

Proof. In the first step, cover O with a grid of almost disjoint cubes of sidelength 1. There are three possibilities;

1. If $C \subset O$ we accept it.
2. If $C \subset O^{\complement}$, we reject them.
3. If $X \cap O \neq \emptyset$ and $C \cap O^{\complement} \neq \emptyset$, we tentatively accept and move to the next step.

In the second step, we make a dyadic decomposition (create 2^n subcubes) of the cubes in O by cutting sidelength in half. We only need do this for the cubes which contained the boundary and repeat the procedure in step 1. We iterate the procedure indefinitely. The end result is that we can find almost disjoint cubes $\{C_j\}_{j=1}^{\infty}$ with $O = \bigcup_{j=1}^{\infty} C_j$. Given $x \in O$, we can find a cube of length 2^k around x contained in O. ■

Note

Unlike the case $n = 1$, this decomposition is non-unique.

1.2 The Exterior Measure

Given **any** subset $E \subset \mathbb{R}^n$, we can define the **exterior measure** $m_*(E)$ generalizing the construction above.

Definition 1.8 (Exterior measure)

Let $E \subset \bigcup_{j=1}^{\infty} C_j$ for cubes C_j. Then

$$m_*(E) := \inf_{C_j} \sum_{j=1}^{\infty} |C_j| \tag{1.1}$$

i.e. take infimum over all coverings of E by cubes.

Note

We have $0 \leq m_*(E) \leq \infty$

Example 1.1

1. $m_*(\{\text{point}\}) = 0$.

   Proof. Let $\{\text{point}\} = \{0\}$, with $\{0\} = \bigcap_{j=1}^{\infty} C_j$ where C_j is a cube centered at zero of length 2^{-j}. Since $\{0\} \subset C_k$ for any k and $m_*(C_k) = 2^{-kn} \to 0$ as $k \to \infty$. ■

2. Let C be a cube. Then $m_*(C) = |C|$.

   Proof. The cube is a covering of itself. Since any other covering yields bigger or equal volume, the infimum is found taking the covering by the cube itself. ■

3. The exterior measure of $\mathbb{R}^n$, $m_*(\mathbb{R}^n)$ is infinite.

   Proof. Take any cube inside $\mathbb{R}^n$ and increase the lengthside. Left as an exercise. ■

4. **Cantor set**. Start with the interval C_0 the unit interval. Let $C_1 = (0, 1/3) \cup (2/3, 1)$. At each step k, remove the middle third of the removing intervals in step $k-1$. At stage C_k, we have 2^k disjoint intervals of length 3^{-k}. The Cantor set $\mathcal{C} = \bigcap_{k=1}^{\infty} C_k$ is uncountable (bijective with $\mathbb{R}$, has cardinality of the continuum), but it is *very* small in the sense of volume (exterior measure); it has measure zero. We know

$$\begin{aligned} m_*(\mathcal{C}) &= m_*\left(\bigcap_{k=0}^{\infty} C_k\right) \\ &\leq m_*(C_N) = \left(\frac{2}{3}\right)^N \xrightarrow{N\to\infty} 0 \end{aligned}$$

   which imply that $m_*(\mathcal{C}) = 0$.

1.3 General properties of exterior measure

1. **Countable sub-additivity**. Given $E = \bigcup_{j=1}^{\infty} E_j \subset \mathbb{R}^n$,

$$m_*(E) \leq \sum_{j=1}^{\infty} m_*(E_j)$$

Proof. Without loss of generality (WLOG), assume $m_*(E_j) < \infty \; \forall \, j$, otherwise we are done. Cover E_j with cubes $\{Q_{k_j}\}$, which imply $E_j \subset \bigcup_{k=1}^{\infty} Q_{k_j}$ with

$$\sum_{k=1}^{\infty} |Q_{k_j}| \leq m_*(E_j) + \frac{\varepsilon}{2^j}$$

for any $\varepsilon > 0$. Clearly,

$$E \subset \bigcup_{j,k=1}^{\infty} Q_{k_j}$$

thus

$$\begin{aligned} m_*(E) &\leq \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} |Q_{k_j}| \\ &\leq \sum_{j=1}^{\infty} \left(m_*(E_j) + \frac{\varepsilon}{2^j} \right) \\ &= \sum_{j=1}^{\infty} m_*(E_j) + \varepsilon. \end{aligned}$$

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

The exterior measure defined on an arbitrary set is too ambitious.

2. If $E_1 \subset E_2$, then $m_*(E_1) \leq m_*(E_2)$

Proof. Any covering of E_2 is a covering of E_1. ■

The following is very useful.

Proposition 1.9

Let $E \subset \mathbb{R}^n$ be any set. Then

$$m_*(E) = \inf_{O} m_*(O)$$

where the infimum is taken over all open cover $O \supset E$.

1.4 Measure

Given any $E \subset \mathbb{R}^n$, we have defined **exterior measure** $m_*(E) = \inf_{Q_j} \sum_{j=1}^{\infty} |Q_j|$ where $E \subset \bigcup_{j=1}^{\infty} Q_j$ covering by cubes. The problem is that m_* is not countably additive for arbitrary disjoint sets. We need to refine the admissible "measurable" subsets of $\mathbb{R}^n$.

Definition 1.10 (Lebesgue measurable)

A subset $E \subset \mathbb{R}^n$ is said to be **Lebesgue measurable** (measurable), written $E \in \mathcal{M}$, provided for every $\varepsilon > 0$, there exists an open set $O \in \mathbb{R}^n$ open with $E \subset O$ such that

$$m_*(O - E) \leq \varepsilon$$

The key point is that we show that $\mathcal{M}$ is closed under countable unions, intersections and taking complements. We call $\mathcal{M}$ a $\sigma-$algebra. Also, given a disjoint countable union $E = \bigcup_{j=1}^{\infty} E_j$, for $E_j \in \mathcal{M}$ and $E_j \cap E_k = \emptyset$ for $j \neq k$, we have

$$m_*(E) = \sum_{j=1}^{\infty} m_*(E_j)$$

Remark

When $E \in \mathcal{M}$, we define Lebesgue measure $m(E) \coloneqq m_*(E)$

We can take about more general measures; take a integer lattice, $\mathbb{Z}^n$. In this discrete infinite set, we can talk about subsets $A \subset \mathbb{Z}^n$, we consider sides rather than the continuum. We could define for instance

$$\mu(k_1, \ldots, k_n) \coloneqq e^{-\phi(k_1,\ldots,k_n)}\delta(\boldsymbol{x} - (k_1, \ldots, k_n)$$

typically called Gibb's measure so that

$$\mu(f) = \sum_k f(k_1, \ldots, k_n)e^{-\phi(k_1,\ldots,k_n)}$$

for $\phi > 0$. We could look at random paths and percolations and look at the limiting object (scaling limit).

Several properties we have to check first:

1. Open set $O \subset \mathbb{R}^n$ is measurable

   Proof. Immediate. Choose $E = O$ in this case. ■

2. Assume $m_*(E) = 0$. Then E is also measurable (*e.g.* $E = \mathcal{C}$, the Cantor set)

   Proof. By last day, we can find an open $O \supset E$ with $m_*(O) \leq m_*(E) + \varepsilon$ for all ε. In this case, $m_*(E) = 0$, therefore $m_*(O) \leq \varepsilon$. Note $O - E \subset O$ so by monotonicity,

$$m_*(O - E) \leq m_*(O) \leq \varepsilon.$$

   Since ε is arbitrary, we are done. ■

3. If $E_j \in \mathcal{M}$ for $j = 1, \dots,$ then $E = \bigcup_{j=1}^{\infty} E_j \in \mathcal{M}$.

   Proof. For each E_j, we choose $O_j \supset E_j$ open with $m_*(O_j - E_j) \leq \frac{\varepsilon}{2^j}$ for all $j = 1, 2, \dots$

   Let $O = \bigcup_{j=1}^{\infty} O_j$ which is open. Clearly, $E = \bigcup_{j=1}^{\infty} E_j \subset O$ and

$$\begin{aligned} m_*(O - E) &\leq \sum_{j=1}^{\infty} m_*(O_j - E_j) \\ &\leq \sum_{j=1}^{\infty} \frac{\varepsilon}{2^j} = \varepsilon \end{aligned}$$

   ■

4. $F \subset \mathbb{R}^n$ closed is measurable.

   Proof. First, it is enough to assume that F is closed and bounded, *i.e* compact. Then, in particular, $m_*(F) < \infty$. Indeed, we can write,

$$F = \bigcup_{k=1}^{\infty} (F \cap B_k)$$

   where B_k is the closed ball of radius $k \in \mathbb{Z}^+$. If we can prove that $F \cap B_k \in \mathcal{M}$, then by the previous proposition, $F \in \mathcal{M}$. Let $K = F \cap B_k$ (compact). We need to find O open, for any $\varepsilon > 0$ $O \supset K$ with $m_*(O - K) \leq \varepsilon$. By open set characterization of exterior measure, for any $\varepsilon > 0$, we can find $O \supset K$ with

$$m_*(O) \leq m_*(K) + \varepsilon \tag{1.2}$$

Note

$O-K$ is open, so by previous result, we can write $O-K=\bigcup_{j=1}^{\infty} C_j$ almost disjoint union of closed cubes. For any N, we let $L=\bigcup_{j=1}^{N} C_j$ (closed) and $L\cap K=\emptyset$ with $d(L,K)=\inf_{x\in L, y\in K}|x-y|$. The proof is left as an exercise.

Since $K\cup L\subset O$, then

$$\begin{aligned} m_*(O) &\geq m_*(K)+m_*(L) \qquad (\text{as } \mathrm{d}(K,L)>0) \\ &= m_*(K)+\sum_{j=1}^{N} m_*(C_j) \end{aligned}$$

which in turns implies

$$\sum_{j=1}^{N} m_*(C_j) \leq m_*(O)-m_*(K) \leq \varepsilon \tag{1.3}$$

Since this is true for any $N\geq 1$, we take the limit as $N\to\infty$ in (1.3) to finally obtain $m_*(O-K)\leq\varepsilon$. ■

5. Complements are measurable, that is given $E\in\mathcal{M}$, then $E^{\complement}\in\mathcal{M}$.

   Proof. Left as an exercise ■

6. Countable intersections of measurable sets are measurable.

   Proof. Write
$$\bigcap_{j=1}^{\infty} E_j = \left(\bigcup_{j=1}^{\infty} E_j^{\complement}\right)^{\complement}$$
   for $E_j\in\mathcal{M}$ for $j=1,\ldots$. Then, use property 5 and closure under countable unions. ■

Theorem 1.11

Suppose $E_1, E_2, \ldots$ are a countable collection of measurable disjoint subsets of $\mathbb{R}^n$. Then $E=\bigcup_{j=1}^{\infty} E_j\in\mathcal{M}$, the measurable sets and

$$m(E)=\sum_{j=1}^{\infty} m(E_j). \tag{1.4}$$

Proof. Approximate E by simple sets that have the countable additivity property in (1.4) and then take limits. Let's assume to start that $m(E_j) < \infty \; \forall \; j = 1, 2, \ldots$

First, we note that (this doesn't require finite measure) by monotonicity of exterior measure,

$$m(E) = m\left(\bigcup_{j=1}^{\infty} E_j\right) \leq \sum_{j=1}^{\infty} m(E_j) \tag{1.5}$$

We need to prove the other direction, that is $m(E) \geq \sum_{j=1}^{\infty} m(E_j)$. There exists $\{F_j\}_{j=1}^{\infty}$ closed (and bounded) with the property that

$$m_*(E_j - F_j) \leq \frac{\varepsilon}{2^j} \qquad \text{for any } \varepsilon > 0$$

(why?)

Exercise 1.1

Consider $E \in \mathcal{M}$ and $O \supset E$ open with $m(O - E) < \varepsilon$. Then apply measurability to $E^{\complement}$ and $\tilde{O}$ with $m(\tilde{O} - E^{\complement}) < \varepsilon$ and take complements.

Choose $N < \infty$ finite. Then $F_1, \ldots, F_n$ are compact (since they are closed and bounded). We already know that

$$m\left(\bigcup_{j=1}^{N} F_j\right) = \sum_{j=1}^{N} m(F_j).$$

But $\bigcup_{j=1}^{N} F_j \subset E$, which implies that

$$\begin{aligned} m(E) &\geq \sum_{j=1}^{N} m(F_j) \\ &\geq \sum_{j=1}^{N} m(E_j) - \varepsilon. \end{aligned} \tag{1.6}$$

for all $\varepsilon > 0$. Now, just take $N \to \infty$ in (1.6) to get

$$m(E) \geq \sum_{j=1}^{\infty} m(E_j).$$

When the E_j's are unbounded, we argue as follows: let $\{Q_j\}_{j=1}^{\infty}$ be closed cubes with $Q_k \subset Q_{k+1}$ and $\bigcup_k Q_k = \mathbb{R}^n$ (an increasing sequence of nested closed cubes). Let $S_1 = Q_1, \ldots, S_k = Q_k - Q_{k-1}$ and let $E_{jk} \coloneqq E_j \cap S_k$. The collection of sets $\{E_{jk}\}_{j,k=1}^{\infty}$ are disjoint and measurable and **bounded**. Clearly, $E_j = \bigcup_{k=1}^{\infty} E_{jk}$ and by the first part, we have that

$$m(E_j) = \sum_{k=1}^{\infty} m(E_{jk}).$$

Now, $E = \bigsqcup_{j=1}^{\infty} E_j = \bigcup_{j,k=1}^{\infty} E_{jk}$, a countable union of disjoint measurable sets. Here $\bigsqcup$ denotes a disjoint union.

So by the first part,

$$m(E) = \sum_{j,k=1}^{\infty} m(E_{jk}) = \sum_{j=1}^{\infty} m(E_j)$$

by (1.6).

■

1.5 Monotone limits of measures

Let $E_1, E_2, \ldots$ be measurable and nested sets, then

Notation

1. when $E = \bigcap_{k=1}^{\infty} E_k$ with $E_k \supset E_{k+1}$, we say that E_k is a decreasing sequence and we write $E_k \searrow E$.

2. when $E = \bigcup_{k=1}^{\infty} E_k$ with $E_k \subset E_{k+1}$, we write $E_k \nearrow E$.

Corollary 1.12

1. If $E_k \nearrow$ E, then $m(E) = \lim_{N\to\infty} m(E_N)$

2. If $E_k \searrow E$ and $m(E_k) < \infty$ for some k, then $m(E) = \lim_{N\to\infty} m(E_n)$.

Proof.

1. We want to use Theorem (1.11) by constructing appropriate countable disjoint union. Let $G_1 = E_1, G_2 = E_2 - E_1, \ldots, G_k = E_k - E_{k-1}$. The G_k's are measurable and disjoint (being the difference of two measurable sets). This imply by Theorem (1.11)

that

$$
\begin{aligned}
m(E) &= \sum_{k=1}^{\infty} m(G_k) \\
&= \lim_{N\to\infty} \sum_{k=1}^{N} m(E_k) \\
&= \lim_{N\to\infty} m\left(\bigsqcup_{k=1}^{N} G_k\right).
\end{aligned}
$$

But $E_N = \bigsqcup_{k=1}^{N} G_k$, therefore $\lim_{N\to\infty} m(E_n) = m(E)$.

2. Left as an exercise.

■

1.6 σ-algebra and Borel sets

We begin with a provisional definition for the purpose of Lebesgue measure, which we generalize in a moment.

Definition 1.13 (σ-algebra (in case of Lebesgue measure))
$\mathcal{A}$ is a collection of Lebesgue measurable subsets of $\mathbb{R}^n$ closed under countable unions, intersections and complements.

The set $\mathcal{M} = \{$ Lebesgue measurable sets$\}$ is a $\sigma-$algebra. The set $\mathbb{B} = \{\sigma$-algebra of Borel sets$\}$, namely the smallest $\sigma-$algebra containing open sets. We have the proper inclusion that $\mathbb{B} \subsetneq \mathcal{M}$ and one can show that $\mathcal{M}$ is the completion of $\mathbb{B}$ (one adds measure zero sets to $\mathbb{B}$).

$\mathbb{B}$ turns ut to be important and so we give the elements of $\mathbb{B}$ names.

Definition 1.14 (G_δ and F_σ sets)
We denote by

$$
\begin{aligned}
G_\delta &= \{ \text{ the countable intersection of open sets}\} \\
F_\sigma &= \{ \text{ the countable unions of closed sets}\}
\end{aligned}
$$

General measure theory

We have

1. $\mathcal{X}$ is a measure space (for example, $\mathcal{X} = \mathbb{R}^n$)
2. a σ-algebra $\mathcal{M}$ is a collection of subsets of $\mathcal{X}$ closed under countable unions, intersections and complementation.
3. a measure $\mu : \mathcal{M} \to [0, \infty]$ with the countable additivity property $\mu\left(\bigsqcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \mu(E_j)$. We denote the data by $(\mathcal{X}, \mathcal{M}, \mu)$.

1.7 Measurable functions

We want to do analysis with measurable sets and **measurable functions**. We need to be able to integrate and then differentiate as well.

The Riemann integral

Let $\{R_k\}_{k=1}^N$ be an almost disjoint collection of closed cubes. Consider the step functions

$$\sum_{k=1}^{N} a_k \chi_{R_k}(x) = S_{N,f}(x)$$

Roughly,

$$\int f(x)\mathrm{d}x = \lim_{N\to\infty} \sum_{k=1}^{N} a_{k,N}|R_k|.$$

This motivates the definition of a measurable function

Definition 1.15 (Characteristic function and simple function)

Given $E \in \mathcal{M}$, we define the **characteristic function** of E

$$\chi_E(x) = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \in E^{\complement}. \end{cases}$$

A **simple function** is of the form $\sum_{k=1}^{N} a_k \chi_{E_k}$ where $E_k \in \mathcal{M}$ for $k = 1, \ldots, N$.

Definition 1.16 (Measurable function)

Consider $f : E \to [-\infty, \infty]$ where $E \in \mathcal{M}$. Then, f is said to be measurable (written $f \in \mathcal{M}_{\mathfrak{e}}$) if $f^{-1}([-\infty, a)) \in \mathcal{M}$ for all $a \in \mathbb{R}$.

Example 1.2

Consider the simple functions

$$S_N(x) = \sum_{i=1}^{N} a_k \chi_{E_k}(x)$$

where $E_k \in \mathcal{M}, x \in \mathbb{R}^n, a \in \mathbb{C}$. Then $S_N(\mathbb{R}^n)$ is a finite set in $\mathbb{R}$. You check that $S_N^{-1}([-\infty, a))$ is measurable. For example, consider $S_N(x) = \chi_{E_k}(x)$.

Measurability is very robust.

Proposition 1.17

We have the following properties

0. The condition $\{f < a\} \in \mathcal{M}$ can be replaced by $\{f \leq a\} \in \mathcal{M}$ or also $\{f > a\} \in \mathcal{M}, \{f \geq a\} \in \mathcal{M}$.

   Proof. To see that $\{f \leq a\} \in \mathcal{M}$ is equivalent to $\{f < a\} \in \mathcal{M}$,

   $$(\Rightarrow) \qquad \{f \leq a\} = \bigcap_{k=1}^{\infty} \left\{ f < a + \frac{1}{k} \right\}$$

   all of which are in $\mathcal{M}$ for all $k \in \mathbb{Z}^+$. Thus, $\{f \leq a\} \in \mathcal{M}$. But also

   $$(\Leftarrow) \qquad \{f < a\} = \bigcup_{k=1}^{\infty} \left\{ f \leq a - \frac{1}{k} \right\}$$

   is in $\mathcal{M}$. The cases $\{f > a\}$ and $\{f \geq a\}$ are treated by taking complements (left as an exercise). ■

1. Suppose $f : \mathbb{R}^n \to (-\infty, \infty)$. Then $f \in \mathcal{M}e$ if and only if $f^{-1}(O) \in \mathcal{M}$ for all $O \subset \mathbb{R}$ for O open.

   Proof. Write $O = \bigcup_{j=1}^{\infty} I_j$ an almost disjoint union of intervals (exercise). For finite-valued functions, it is enough to check that $I = \{a < f < b\} \in \mathcal{M}$ for any $a, b \in \mathbb{R}$. ■

   Remark

   Similarly, one can show that $f \in \mathcal{M}e$ if and only if $f^{-1}(F) \in \mathcal{M}$ for all $F \in \mathbb{R}$, Fclosed

2. (a) If $f \in \mathcal{C}^0(\mathbb{R}^n)$, then $f \in \mathcal{M}e(\mathbb{R}^n)$

   (b) If $f \in \mathcal{M}e(\mathbb{R}^n)$ and finite-valued and $\Phi \in \mathcal{C}^0(\mathbb{R}^n)$, then $\Phi \circ f \in \mathcal{M}e(\mathbb{R}^n)$.

   Proof.

   (a) $f \in \mathcal{C}^0 \Rightarrow f^{-1}((-\infty, a)) = O$ open and so $O \in \mathcal{M}$.

   (b) $\Phi^{-1}((-\infty, a)) = O$ is open, and

   $$(\Phi \circ f)^{-1}((-\infty, a)) = f^{-1}(O) \in \mathcal{M}$$

■

Remark

$f \circ \Phi \notin \mathcal{M}e$ in general.

We have to show that given $\{f_n\}_{n=1}^{\infty}$ with $f_n \in \mathcal{M}e$, we want to show that inf, sup and limits (when they exist) are all measurable.

3. Suppose $\{f_n\}_{n=1}^{\infty} \in \mathcal{M}e$. Then
   (a) $\sup_n f_n \in \mathcal{M}e$
   (b) $\inf_n f_n \in \mathcal{M}e$
   (c) $\limsup_{n\to\infty} f_n \in \mathcal{M}e$
   (d) $\liminf_{n\to\infty} f_n \in \mathcal{M}e$

Proof.

(a) $\{\sup_n f_n > a\} = \bigcup_{n=1}^{\infty}\{f_n > a\} \in \mathcal{M}e$

(b) Use that $\inf_n f_n(x) = -\sup_n(-f_n(x))$ and then use (3a)

(c) $\limsup f_n = \inf_n \sup_{k\geq n} f_k$. By (3a), $\sup_{k\geq n} f_k \in \mathcal{M}e$ and $\inf_k(\cdots) \in \mathcal{M}e$ by (3b)

(d) $\liminf f_n = \sup_n\{\inf_{k\geq n} f_k\} \in \mathcal{M}$ by (3a) and (3b).

Corollary 1.18

Assume that $f_n : \mathbb{R}^n \to (-\infty, \infty)$ are all measurable, $n = 1, 2, \ldots$ and that $\lim_{n\to\infty} f_n = f$ exists pointwise (or a.e.). Then, $f \in \mathcal{M}e(\mathbb{R}^n)$. This is crucial.

■

Section 2
Lebesgue integral

We defined simple functions

$$s(x) = \sum_{i=1}^{N} c_i \chi_{E_j}(x) \qquad c_i > 0$$

where $E_j \in \mathcal{M}$ and χ_E is the indicator function. Step functions are special cases, where $E_i = Q_i$, where Q_i be cubes in $\mathbb{R}^n$.

Definition 2.1

Given $E \in \mathcal{M}$, we define the **Lebesgue integral** of $s(\boldsymbol{x}) \geq 0$ to be

$$I_E(s) = \int_E sdm := \sum_{i=1}^{n} c_i m(E_i \cap E)$$

Proposition 2.2

1. Given $c \in \mathbb{R}$, $I_E(cs) = cI_E(s)$ (linearity)
2. If $s_1, s_2 : \mathbb{R}^n \to \mathbb{R}^+$ are simple functions, then

$$I_E(s_1 + s_2) = I_E(s_1) + I_E(s_2).$$

Proof. Given $s_1 = \sum_{i=1}^{M} c_i \chi_{E_1}$ and $s_2 = \sum_{j=M+1}^{N} d_j \chi_{F_j}$ and $s_1 + s_2 = \sum_{k=1}^{N} g_k \chi_{G_k}$ and wlog $M \leq N$ where

$$g_k = \begin{cases} c_k & \text{if } 1 \leq k \leq M \\ d_k & \text{if } M+1 \leq k \leq N \end{cases} \qquad G_k = \begin{cases} E_k & \text{if } 1 \leq k \leq M \\ F_k & \text{if } M+1 \leq k \leq N \end{cases}$$

and

$$\begin{aligned} I_E(s_1 + s_2) &= \sum_{k=1}^{N} g_k m(E \cap G_k) \\ &= \sum_{i=1}^{M} c_i m(E \cap E_i) + \sum_{i=M+1}^{N} d_i m(E \cap F_i) \end{aligned}$$

■

Lebesgue integral for non-negative measurable functions

Definition 2.3

Given $f \in \mathcal{M}e(\mathbb{R}^n)$ and $f \geq 0$, for $E \in \mathcal{M}$, the Lebesgue integral of f

$$I_E(f) = \int_E f dm := \sup \{I_E(s); 0 \leq s \leq f\}$$

where the $s(\boldsymbol{x})$ are simple functions and the sup is over the partition.

Proposition 2.4

Given $f = s$, a fixed simple function,

$$\int_E f dm = I_E(s) = \sum_{i=1}^{N} c_i m(E_i \cap E),$$

Proof. Left as an exercise. Clearly, $\sup\{I_E(s)\} \leq \sum c_i m(E_i \cap E)$ and is attained when $f = s$ itself. Choose any other simple function, then show that it will be strictly less as they are below. ■

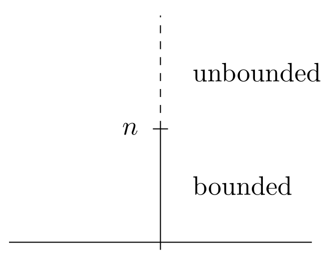
Theorem 2.5

Given any non-negative measurable function $f : \mathbb{R}^n \to [0, \infty]$ there exists a non-negative monotone sequence of simple functions $0 \leq s_1 \leq s_2 \leq \dots$ and $s_N \leq f$ such that

$$\lim_{N\to\infty} s_N(\boldsymbol{x}) = f(\boldsymbol{x}) \tag{2.7}$$

for all $\boldsymbol{x} \in \mathbb{R}^n$ (written $s_N \nearrow f$). Moreover, if $f(\boldsymbol{x})$ is bounded *i.e.*($|f(\boldsymbol{x})| \leq M \leq \infty \; \forall \; \boldsymbol{x} \in \mathbb{R}^n$), then the convergence in (2.7) is uniform.

Proof. We write $[0, \infty] = [0, n) \cup [n, \infty]$ for some $n > 0$.



Bounded piece: Write $[0, n)$ as a disjoint union of intervals

$$I_i = \left\{ t \in \mathbb{R}; \frac{i-1}{2^n} \leq t \leq \frac{i}{2^n} \right\}$$

for $i = 1, 2, \dots, n2^n$. Let $E_i = f^{-1}(I_i) \in \mathcal{M}$ since $f \in \mathcal{M}_{\boldsymbol{e}}$,

$$F_n = f^{-1}([n, \infty]) \in \mathcal{M}$$

Clearly,

$$\mathbb{R}^n = \bigsqcup_{i=1}^{n2^n} E_i \bigsqcup F_n$$

a mutually disjoint decomposition. Define

$$s_n(\boldsymbol{x}) = \sum_{i=1}^{n2^n} \left(\frac{i-1}{2^n} \right) \chi_{E_i}(\boldsymbol{x}) + n\chi_{F_n}(\boldsymbol{x})$$

for $E_i = f^{-1}(I_i)$.

Note

When $\boldsymbol{x} \in E_i$, $s_n(\boldsymbol{x}) = (i-1)/2^n \leq f(\boldsymbol{x})$ since $(i-1)/2^n \leq f \leq i/2^n$, therefore $s_n(\boldsymbol{x}) \leq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in f^{-1}([0, n))$

When $\boldsymbol{x} \in F_n$, $f(\boldsymbol{x}) \geq n = s_n(\boldsymbol{x})$, therefore $s_n(\boldsymbol{x}) \leq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{R}^n$.

Claim

We have $\lim_{n\to\infty} s_n(\boldsymbol{x}) = f(\boldsymbol{x})$ In the first case, $f(\boldsymbol{x}) = \infty \Rightarrow \boldsymbol{x} \in F_n$ for all n, so $s_n(\boldsymbol{x}) \to \infty$.

In the second case, suppose $f(\boldsymbol{x}) < n_0 < \infty$. Then, for $n > n_0$

$$\frac{i-1}{2^n} < f(\boldsymbol{x}) < \frac{i}{2^n}$$

for some i. Then $s_n(\boldsymbol{x}) = (i-1)/2^n$ imply $|f(\boldsymbol{x}) - s_n(\boldsymbol{x})| < \frac{1}{2^n}$ as $n \to \infty$.

Exercise 2.1

Show that this convergence is uniform provided f is bounded (immediate in fact).

■

Given a measurable set $E \in \mathcal{M}$ and non negative functions $f, g \geq 0$ and measurable, then $\int_E (f+g) dm = \int_E f cm + \int_E g dm$ is clear, but surprinsingly is a bit tricky to prove.

Given $f \geq 0, f \in \mathcal{M}e(\mathbb{R}^n)$, we constructed the Lebesgue integral

$$I_E(f) = \int_E f dm = \sup\{I_E(s) : 0 \leq s \leq f, s \text{ simple }\}$$

and we proved some basic things for $I_E(f)$, notably

1. $c \int_E s = \int_E cs$
2. $\int_E (s_1 + s_2) = \int_E s_1 + \int_E s_2$ for $s_1, s_2, \geq 0$ simple functions
3. If $E \subset F, E, F \in \mathcal{M}$ and $s \geq 0$ is simple

$$\int_E s \leq \int_F s$$

4. If $0 \leq s_1 \leq s_2$ are simple functions and $E \in \mathcal{M}$, then $\int_E s_1 \leq \int_E s_2$

We try to extend to general non-negative measurable functions these facts.

Proposition 2.6

In the following we take $E, \in \mathcal{M}$ to be measurable sets and $f, g \geq 0$ measurable functions. Then

1. Assume $f \leq g$, then $\int_E f \leq \int_E g$.
2. If $E \subset F$, then $\int_E f \leq \int_F f$
3. If $m(E) = 0$, then $\int_E f dm = \int_E f = 0$.

Proof.

1. If $0 \leq s \leq f$ where as usual $s(x)$ is a simple function, then also $0 \leq s \leq g$, thus

$$\sup\{I_E(s); 0 \leq s \leq f\} \leq \sup\{I_E(s); 0 \leq s \leq g\}$$

   and by definition, this holds if and only if $\int_E f dm \leq \int_E g dm$.

2. Check first for $f = \chi_G$, where $G \in \mathcal{M}$. Clearly, $\int_E \chi_G dm = m(E \cap G)$ where $E \in \mathcal{M}$. Also, $\int_F \chi_G dm = m(G \cap F)$, and since $E \subseteq F$, which implies that $(E \cap G) \subseteq (F \cap G)$. By monotonicity of measure, $m(E \cap G) \leq m(F \cap G)$. The result is true for simple functions (why?) and thus true for all non-negative function $f \in \mathcal{M}e$.

3. Suppose $f(x) = s(x) \geq 0$ simple, with $s(x) = \sum_{i=1}^n c_i \chi_{E_i}(x)$. Suppose $m(E) = 0$, then $I_E(s) = \sum_{i=1}^n c_i m(E \cap E_i)$ and since $m(E) = 0$ and $0 \leq m(E \cap E_i) \leq m(E)$ by monotonicity, this imply $m(E \cap E_i) = 0$ for all i and as a result $I_E(s) = 0$ for all $s \geq 0$ simple and $\int_E f = 0$.

We leave the linearity property $\int_E (f+g)dm = \int_E fdm + \int_E gdm$. ■

Theorem 2.7 (Chebyshev inequality)
Suppose $f \geq 0$ and $f \in \mathcal{M}e(\mathbb{R}^n)$ and $E \in \mathcal{M}$. Then, if $c > 0$,

$$m(\{x \in E | f(x) \geq c\}) \leq \frac{1}{c}\left(\int_E fdm\right)$$

and this can be extended to L_p inequalities.

Proof. Let E_c=$\{x \in E | f(x) \geq c\}$ since $f \geq c$ on E_c, this imply

$$\int_{E_c} fdm \geq c\int_{E_c} dm = cm(E_c) \tag{2.8}$$

by property 1. By property 2, we can enlarge the left hand side by taking the integral over $E_c \subseteq E$,

$$\int_E fdm \geq \int_{E_c} fdm. \tag{2.9}$$

So by (2.8) and (2.9), we have

$$\int_E fdm \geq cm(E_c)$$

■

Corollary 2.8
Suppose $f \geq 0$ and $f \in \mathcal{M}e(\mathbb{R}^n), E \in \mathcal{M}$ with $\int_E fdm < \infty$. Then $m(\{x \in E : f(x) = \infty\}) = 0$, *i.e.* measurable functions with finite integral can't be too bad.

Proof. Define $A_n = \{x \in E : f(x) \geq n\}$, $n \in \mathbb{Z}^+$ and $A = \{x \in E : f(x) = \infty\}$. Clearly, both sets are measurable and $A \subset \cap_{n=1}^{\infty} A_n \subset A_N$ for some $N \in \mathbb{Z}^+$. This imply $m(A) \leq m(A_N)$ by monotonicity for any N. The right hand side is by Chebyshev less than or equal to

$$m(A_N) \leq \frac{1}{N}\left(\int_E fdm\right) \tag{2.10}$$

and the integral is finite by assumption. Since N is arbitrary, letting $N \to \infty$ in (2.10) imply $m(A) = 0$. ■

Corollary 2.9

Suppose $f \geq 0, f \in \mathcal{M}e$, $E \in \mathcal{M}$. If $\int_E f dm = 0$, then $f = 0$ almost everywhere on E.

Note

We say that $f = 0$ a.e. on $E \in \mathcal{M}$ provided that $m(\{x \in E : f(x) \neq 0\}) = 0$.

Proof. Let $A = \{x \in E : f(x) \neq 0\}$ and $A_n = \{x \in E : f(x) > n^{-1}\}$. The set $A = \bigcup_{n=1}^{\infty} A_n$ and as such, $m(A) \leq \sum_{n=1}^{\infty} m(A_n)$ by subadditivity. We use Chebyshev again to estimate from above the measure of A_n.

$$0 \leq m(A_n) \leq n \left(\int_E f dm \right) = 0$$

for all n, so $m(A_n) = 0 \; \forall \; n = 1, 2, \ldots$ imply $m(A) = 0$. ■

2.1 Probability and measure

We can view probability as being equivalent to measure theory (with bells and whistles). The basic idea is the following; take a probability (sample) space $(\Omega, \mathcal{A}, \mathsf{P})$ is equivalent to X a measurable space with $\mu(X) = 1$ (normalized sequence). An event corresponds to an element of a measurable set. The probability itself of a series of events is the measure $\mu(E) \leq 1$.

Definition 2.10 (Bernoulli sequence)

A Bernoulli sequence is an infinite (fair) coin toss, we can identify this with infinite binary expansions.

$$H \; T \; H \; H \; T \; T \; T \ldots \equiv 1011000 \ldots$$

where $1 \equiv H$ and $0 \equiv T$. The first point is the following

Definition 2.11 (Terminating binary expansion)

We say that a binary expansion $\omega = .a_1 a_2 a_3 \ldots$, where $a_j \in \{0, 1\}$ is **terminating** if $a_n = 0$ for $n \geq N$.

Proposition 2.12

Let A be the set of all binary expansions (Bernoulli trials) and let A_{reg} be the set of all terminating binary expansions. Then,

$$A \setminus A_{\text{reg}} \cong [0, 1]$$

namely there is a bijection between the non-terminating binary expansions and the unit interval.

Proof. Uses Exercise on Homework 1 and the fact that given $\omega \in [0,1]$, we can write $\omega = \sum_{j=1}^{\infty} \frac{a_j}{2^j}$ for $a_j \in \{0,1\}$. The only thing left to show is that non-terminating expansions of a given $\omega \in I = [0,1]$ is **unique** [exercise]. ■

So (ignoring A_{reg}), we can identify {Bernoulli trials} with the unit interval $[0,1]$. Thus probability is Lebesgue measure on $[0,1]$.

Example 2.1

Suppose E is the event that an H occurs in the N^{th} place. To compute probability of an event, we fix the first $N-1$ trials $a_1 a_2 \dots a_{N-1}$

$$s = .a_1 a_2 \dots a_{N-1} 1 a_{N+1} \dots$$

All elements of E with $(a_1, \dots, a_{N-1})$ fixed correspond in term of dyadic expansion to the interval $E_N = \left(s, s + 2^{-N}\right]$. Note $\sum_{j=N+1}^{\infty} 2^{-j} = 2^{-N}$ by randomizing the digits after N^{th} position. So the probability $\mathsf{P}(E) = m((s, s+2^{-N}]) \times 2^{N-1}$ ranging the first $N-1$ digits, which gives $2^{-N} 2^{N-1} = 1/2$.

2.2 Rademacher Functions

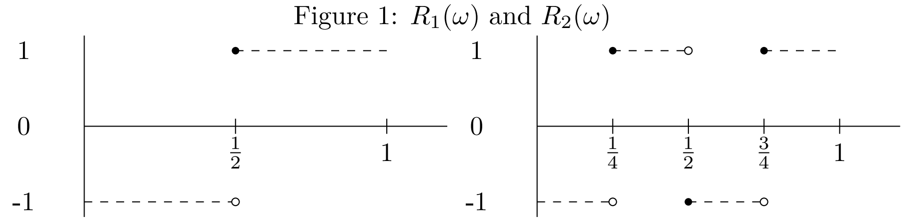
Definition 2.13 (Rademacher functions)

Given $\omega = .a_1 a_2 \dots \in I = [0,1]$, we define the k^{th} Rademacher function to be

$$R_k(\omega) = 2a_k - 1, \qquad R_k : [0,1] \to \{-1, 1\}$$

Lets look at $R_1(\omega)$: Clearly, $R_k \in \mathcal{M}e([0,1])$. It is a step function.

Figure 1: $R_1(\omega)$ and $R_2(\omega)$



To understand important issues like "gambler's ruin", we define the associated functions

$$\begin{aligned} S_N(\omega) &:= \sum_{i=1}^{N} R_k(\omega) \\ &= 2\sum_{j=1}^{N} a_k - N \end{aligned}$$

and $S_n(\omega)$ gives the total amount of money won or lost at the N^{th} stage. For example the problem of gambler's ruin can be phrased in terms of the S_N: given an initial stake of X, we consider the set

$$E_N = \{\omega \in I | S_l(\omega) > -X \text{ for } l \leq N-1 \text{ and } S_N(\omega) = -X\}$$

The gambler's ruin at state N amounts to computing the Lebesgue measure of $E_n, m(E_n)$. E_N is the intersection of finitely-many measurable sets and this is measurable. Another important issue concerns of laws of large numbers, *i.e.* as $N \to \infty$, what is the probability of some number of H and T and the rate of convergence. To study this, we fix $\varepsilon > 0$ and consider the following measurable set

$$E_N = \left\{\omega \in I, \left|\frac{S_N(\omega)}{N} - \frac{1}{2}\right| > \varepsilon\right\}$$

and $\left|\frac{S_N(\omega)}{N} - \frac{1}{2}\right|$ is definitively measurable ($S_N(\omega)$ is a step function, $\frac{1}{2}$ a constant and $|\cdot|$ is also a continuous function).

Theorem 2.14 (Weak law of large number for Bernoulli trials)
For any fixed $\varepsilon > 0$,

$$\lim_{N\to\infty} m(E_N) = 0$$

Proof. We have the equivalent statement

$$\left|\frac{S_N(\omega)}{N} - \frac{1}{2}\right| \quad \Leftrightarrow \quad |2S_N(\omega) - N| > 2N\varepsilon$$

So in terms of $S_N(\omega)$, we have to estimate

$$\begin{aligned}
m(\{\omega \in I : |S_N(\omega)| > N\varepsilon\}) &= m\left(\{\omega \in I; S_N^2(\omega) > 4N^2\varepsilon^2\}\right) \\
&\leq \left(\frac{1}{4N^2\varepsilon^2}\right) \int_I S_N^2(\omega) dm \\
&= \left(\frac{1}{4N^2\varepsilon^2}\right) \int_0^1 S_N^2(\omega) d\omega \\
&= \left(\frac{1}{4N^2\varepsilon^2}\right) \sum_{i=1}^N \int_I R_k^2 dm + \sum_{i \neq j} \int_I R_i R_j dm;
\end{aligned}$$

using Chebyshev's inequality; we will prove shortly the equivalence of Lebesgue and Riemann integral when the two agree. Since $R_K^2(\omega) = 1$ for all k and $\omega \in I$, we have $\int_I R_k^2(\omega) dm = 1$. On the other hand, suppose WLOG that $i < j$ and consider the interval

$$J = \left(\frac{l}{2^i}, \frac{l+1}{2^i}\right], \qquad 0 \leq l \leq 2^i - 1$$

On J, R_i oscillates $2(j-i)$ times implies $\int_J R_i R_j dm = 0$. We thus get

$$\left(\frac{1}{4N^2\varepsilon^2}\right) N = \frac{1}{4N\varepsilon^2}$$

Therefore $m(A_N) \leq (4N\varepsilon^2)^{-1} \to 0$ as $N \to \infty$. ■

This is a small step, but the question we are really interested in is the following. Consider

$$N = \left\{\omega \in I; \lim_{N \to \infty} \left|\frac{S_N(\omega)}{N} - \frac{1}{2}\right| = 0\right\}$$

Theorem 2.15 (Strong law of large numbers for Bernoulli trials)
We have $m(N^{\complement}) = 0$ where $N^{\complement} = I \setminus N$.

Remark
Even though $m(N^{\complement}) = 0$, it is nasty and "very" uncountable. Consider the map $\sigma : I \to I$. $\sigma(\omega) = \omega'$, where $\omega = .a_1 a_2 \ldots$ and $\omega' = .a_1 11 a_2 11 a_3 \ldots$

Clearly, $\sigma(\cdot)$ is bijective, so card(image(σ)) = card[0,1].

Claim
Image$(\sigma) \subset N^{\complement}$

Proof. It is easy to check (exercise) that $S_{2n}(\omega') \geq 2n$ which imply $S_{3n}(\omega')/3n \geq 2n/3n = 2/3$. ■

The proof of the Strong law of large numbers is due to Mark Kac (1963).

Proof. Given $\varepsilon > 0$, consider

$$A_N = \left\{\omega \in I; S_N^4(\omega) \geq \varepsilon^4 N^4\right\}$$

We apply Chebyshev as before:

$$m(A_N) \leq \left(\frac{1}{\varepsilon^4 N^4}\right)\left(\int_I S_N^4(\omega)dm\right). \tag{2.11}$$

Writing out $\int_I S_N^4(\omega)dm$, we get terms of the following type:

$$\int_I R_\alpha^4, \qquad \alpha = 1, \dots, N \tag{2.12a}$$

$$\int_I R_\alpha^2 R_\beta^2 \qquad \alpha \neq \beta \tag{2.12b}$$

$$\int_I R_\alpha^2 R_\beta R_\gamma, \qquad \alpha \neq \beta \neq \gamma \tag{2.12c}$$

$$\int_I R_\alpha^3 R_\beta, \qquad \alpha \neq \beta \tag{2.12d}$$

$$\int_I R_\alpha R_\beta R_\gamma R_\delta, \qquad \alpha \neq \beta \neq \gamma \neq \delta \tag{2.12e}$$

Since $R_\alpha^4 = 1$ and $R_\alpha^2 R_\beta^2 = 1$, the first two terms give $\int_I R_\alpha^4 = \int_I R_\alpha^2 R_\beta^2 = 1$. Because of the oscillation property, these two terms (namely (2.12a) (2.12b)) are the only ones contributing. Consider $J = \left(\frac{l}{2^j}, \frac{l+1}{2^j}\right]$ and wlog assume $\alpha < \beta < \gamma < \delta$, which imply $\int_J R_\alpha R_\beta R_\gamma R_\delta = 0$ by relabelling indices in (2.12e). This imply

$$m(A_N) \leq \frac{1}{N^4\varepsilon^4}\left[3N^2 - 2N\right] \leq \frac{3N^2}{N^4\varepsilon^4}$$

(count the number of the terms in (2.12b)). We now have a convergent series and we can choose ε as a function. So far, we have the improved estimate $m(A_N) = O\left(\frac{1}{N^2\varepsilon^4}\right)$.

We postpone the last step in Strong Law (homework 2). ■

2.3 Convergence theorems

There are three main results in this section, namely the monotone convergence, dominated convergence and Fatou's (and reverse Fatou's) lemma.

Theorem 2.16 (Monotone convergence)

Let $E \in \mathcal{M}$, consider a sequence of measurable functions $\{f_n\}_{k=1}^{\infty} \in \mathcal{M}e$, $f_n \geq 0$ and suppose

$$0 \leq f_1 \leq f_2 \leq \cdots .$$

Let $f = \lim_{n\to\infty} f_n \in \mathcal{M}e$. Then,

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

The idea of the proof is to decompose E into a disjoint union of sets

$$E = \bigcup_{i=1}^{\infty} A_i, \qquad \text{where } A_i \cap A_j = \emptyset \text{ if } i \neq j$$

and then use

$$\int_E f dm = \sum_{i=1}^{\infty} \int_{A_i} f dm \tag{2.13}$$

The first step is to prove (2.13).

Theorem 2.17

Assume that $f \in \mathcal{M}e$ and $f \geq 0$ and suppose $A_1, A_2, \ldots$ are pairwise disjoint measurable sets. Let $A = \bigcup_{i=1}^{\infty} A_i \in \mathcal{M}$. Then

$$\int_A f dm = \sum_{i=1}^{\infty} \int_{A_i} f dm.$$

Proof. First, consider $f = \chi_E, E \in \mathcal{M}$. Then

$$\int_A \chi_E = m(E \cap A)$$
$$\int_{A_i} \chi_E = m(E \cap A_i); \qquad i = 1, 2 \ldots$$

which imply

$$E \cap A = \bigcup_{i=1}^{\infty} (E \cap A_i)$$

where $E \cap A_i$ are pairwise-disjoint. This implies

$$\begin{aligned} m(E \cap A) &= \sum_{i=1}^{\infty} m(E \cap A_i) \\ \Leftrightarrow \quad \int_A \chi_E &= \sum_{i=1}^{\infty} \int_{A_i} \chi_E \end{aligned}$$

By linearity, we have that for any $0 \leq s(\boldsymbol{x})$ simple,

$$\int_A sdm = \sum_{i=1}^{\infty} \int_{A_i} sdm \tag{2.14}$$

Now consider the general case, $f \in \mathcal{M}_{\mathcal{E}}$, $f \geq 0$. Given $\varepsilon > 0$, we can find $s(\boldsymbol{x})$ simple with $0 \leq s \leq f$ such that

$$\int_A fdm \leq \int_A sdm + \varepsilon$$

By the last step,

$$\int_A sdm = \sum_{i=1}^{\infty} \int_{A_i} sdm$$

which imply

$$\begin{aligned} \int_A f(dm) &\leq \sum_{i=1}^{\infty} \int_{A_i} sdm + \varepsilon \\ &\leq \sum_{i=1}^{\infty} \int_{A_i} fdm + \varepsilon. \end{aligned}$$

since $0 \leq s \leq f$. Since $\varepsilon > 0$ is arbitrary, we have

$$\int_A fdm \leq \sum_{i=1}^{\infty} \int_{A_i} fdm.$$

For the opposite direction, we want $\int_A f \geq \sum_{i=1}^{\infty} \int_{A_i} f$. We approximate from below. Assume for the moment that $A = A_1 \cup A_2, A_1 \cap A_2 = \emptyset$ where $A_i \in \mathcal{M}$. Given $\varepsilon > 0$, we can find $0 \leq s_1 \leq f$ with

$$\int_{A_1} s_1 \geq \int_{A_1} f - \frac{\varepsilon}{2} \tag{2.15}$$

Similarly, over A_2, we can find $0 \le s_2 \le f$ with

$$\int_{A_2} s_2 \ge \int_{A_2} f - \frac{\varepsilon}{2}$$

Note that $s \coloneqq \max(s_1, s_2)$ is also simple with $0 \le s \le f$ and both identities (2.14) and (2.15) hold. Thus

$$\int_{A_i} s \ge \int_{A_i} f - \frac{\varepsilon}{2} \tag{2.16}$$

with

$$\int_A s = \int_{A_1} s + \int_{A_2} s$$

since $s(\boldsymbol{x})$ is simple. By (2.16),

$$\int_A s \ge \int_{A_1} f = \int_{A_2} f - \varepsilon$$

for all $\varepsilon > 0$ and since $0 \le s \le f$, by monotonicity,

$$\int_A f \ge \int_A s \ge \sum_{i=1}^{2} \int_{A_i} f - \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary,

$$\int_A f \ge \sum_{i=1}^{2} \int_{A_i} f.$$

The same argument works if $A = \bigcup_{i=1}^n A_i$, $A_i \in \mathcal{M}$ parwise disjoint. This mean

$$\int_A f \ge \sum_{i=1}^{n} \int_{A_i} f$$

for any $n \ge 1$ finite. In the general case, $A = \bigsqcup_{i=1}^{\infty} A_i$ ans since $A_1 \cup \cdots \cup A_n \subset A$, by monotonicity

$$\int_A f \ge \int_{\sqcup_{i=1}^n A_i} f = \sum_{i=1}^{n} \int_{A_i} f \tag{2.17}$$

by the last step. Since n is arbitrary, we take limits as $n \to \infty$ in (2.17) and we finally get

$$\int_A f \geq \sum_{i=1}^{\infty} \int_{A_i} f.$$

■

We now continue with a lemma for the proof of the **monotone convergence theorem.** We first start with a direct consequence of the above theorem.

Lemma 2.18

Let $f \geq \in \mathcal{M}e$ and $E_1, E_2, \ldots \in \mathcal{M}$ nested sets where $E_1 \subset E_2 \subset E_3 \subset \ldots$ and set $E = \bigcup_{i=1}^{\infty} E_i \in \mathcal{M}$. Then

$$\int_E f dm = \lim_{i \to \infty} \int_{E_i} f dm$$

Proof. Set $A_1 = E_1, A_2 = E_1 \setminus E_2, A_3 = E_3 - E_2, \ldots$ Then $E = \bigcup_{i=1}^{\infty} A_i$ and $A_i \cap A_j = \emptyset \; \forall \; i \neq j$ since E_j' are nested. So by the previous theorem

$$\begin{aligned} \int_E f dm &= \sum_{i=1}^{\infty} \int_{A_i} f dm \\ &= \lim_{n \to \infty} \int_{E_n} f dm \end{aligned}$$

since $E_n = \bigcup_{i=1}^{\infty} A_i$. First, since $\lim_{n \to \infty} f_n = f$ and $f_n \leq f$ for all n,

$$\lim_{n \to \infty} \int_E f_n \leq \int_E f$$

by monotonicity. We want to prove the opposite direction, namely

$$\lim_{n \to \infty} \int_E f_n \geq \int_E f$$

there, we decompose $E = \bigcup_{n=1}^{\infty}$ where $E_n \subset E_{n+1}$ and just use $\int_E f = \lim_{n \to \infty} \int_{E_n} f$. Let $s(\boldsymbol{x})$ be simple with $0 \leq s \leq f$ and choose $0 < c < 1$. We define

$$E_n \coloneqq \{x \in E | f_n(\boldsymbol{x}) \geq cs(\boldsymbol{x})\}$$

Then, we have the following

Claim

Let $E = \bigcup_{n=1}^{\infty} E_n$ and note that since $f_n \leq f_{n+1}$, we have that $E_n \subset E_{n+1}$ for all $n \geq 1$ by monotonicity. The proof is left as an exercise

The claim implies

$$\begin{aligned}
\lim_{n\to\infty}\int_{E_n} f_n &\geq \int_E f_n && \forall\ n\\
&\geq \int_{E_n} f_n && \text{(by monotonicity)}\\
&\geq c\int_{E_n} s(\boldsymbol{x})
\end{aligned}$$

The upshot is

$$\lim_{n\to\infty}\int_E f_n \geq c\int_{E_k} s(\boldsymbol{x}) \tag{2.18}$$

for all $k \geq 1$ and $0 < c < 1$. We can take limit over k, $\lim_{k\to\infty}$ in (2.18). By the previous lemma,

$$\lim_{n\to\infty}\int_E f_n \geq c\lim_{k\to\infty}\int_{E_k} s(\boldsymbol{x}) = c\int_E s(\boldsymbol{x})$$

since $E_k \subset E_{k+1}$. Then

$$\begin{aligned}
\lim_{n\to\infty}\int_E f_n &\geq c\int_E s(\boldsymbol{x})\\
&\geq c\int_E f - \varepsilon
\end{aligned}$$

for any $\varepsilon > 0$ choosing $0 \leq s \leq f$; since ε is arbitrary, let $\varepsilon \to 0^+$ and since c is also arbitrary we can take $c \to 1$.[1] Thus

$$\lim_{n\to\infty}\int_E f_n \geq \int_E f$$

■

2.4 Applications of monotone convergence theorem (MCT)

Recal that we postponed the proof of linearity of the Lebesgue integral, we wanted to show the following

[1] The constant c is needed to prove the claim.

Theorem 2.19

Assume $f, g \geq 0$,[2] where $f, g \in \mathcal{M}e$. Given $E \in \mathcal{M}$, then

$$\int_E (f+g)dm = \int_E f dm + \int_E g dm$$

Proof. Given $f \geq 0$, $f \in \mathcal{M}e$, we have constructed a sequence $\{s_n\}_{n=1}^{\infty}$ of simple function with $0 \leq s_n, s_n \nearrow f$. Similarly, $\{s'_n\}_{n=1}^{\infty}$ simple functions with $0 \leq s' \nearrow g$ (so that $0 \leq s_1 \leq s_2 \leq \cdots, \lim_n s_n = f$).

Then $0 \leq s_n + s'_n \nearrow f + g$, therefore by the monotone convergence theorem,

$$\begin{aligned}\lim_{n\to\infty} \int_E (s_n + s'_n)dm &= \int_E \lim_{n\to\infty} (s_n + s'_n)dm \\ &= \int_E (f+g)dm.\end{aligned}$$

But

$$\int_E (s_n + s'_n) = \int_E s_n + \int_E s'_n \tag{2.19}$$

for $n \geq 1$. Since limit of all terms in (2.19) exist by monotonicity, this imply

$$\begin{aligned}\int_E f + g &= \lim_{n\to\infty} \int_E (s_n + s'_n) \\ &= \lim_{n\to\infty} \int_E s_n + \lim_{n\to\infty} \int_E s'_n \\ &= \int_E f + \int_E g\end{aligned}$$

by the monotone convergence theorem. ∎

We also have the following corollary for infinite series

Corollary 2.20

Suppose $\{f_n\}_{n=1}^{\infty}$ is a sequence of **non-negative** measurable functions and let $E \in \mathcal{M}$. Then

$$\sum_{k=1}^{\infty} \int_E f_k dm = \int_E \left(\sum_{k=1}^{\infty} f_k \right) dm.$$

[2] The non-negativity is crucial for the MCT, as the result is false otherwise

Proof. Let $S_n = \sum_{k=1}^n f_k$, $\{S_n\}_{n=1}^\infty$ is a sequence of partial sums. Since $f_k \geq 0, S_n \leq S_{n+1}$ for all $n \geq 1$ (since we have a monotone non-decreasing sequence). Thus

$$\begin{aligned}
\int_E \sum_{k=1}^\infty f_k dm &= \int_E \left(\lim_{n\to\infty} S_n\right) dm \\
&= \lim_{n\to\infty} \int_E S_n dm \\
&= \lim_{n\to\infty} \int_E (f_1 + \cdots + f_n) dm \\
&= \lim_{n\to\infty} \sum_{k=1}^n \int_E f_k dm \\
&= \sum_{k=1}^\infty \int_E f_k dm.
\end{aligned}$$

by MCT and the previous result. ■

The monotone convergence theorem also allows us to extend the definition of the Lebesgue integral to measurable functions of arbitrary sign. We do this by splitting $f \in \mathcal{M}e$ into its positive and negative part.

We do this by splitting $f \in \mathcal{M}e$ into its positive and negative parts

$$\begin{aligned}
0 \leq f_+ &:= \max(f, 0) \in \mathcal{M}e \\
0 \leq f_- &:= \max(-f, 0) \in \mathcal{M}e
\end{aligned}$$

and the key obvious point is that we can write $f = f_+ - f_-$.

The condition that $\int_E |f| < \infty \Leftrightarrow \int_E f_+ < \infty, \int_E f_- < \infty$ for $E \in \mathcal{M}e$. This is left as an exercise.

Definition 2.21 ($\mathcal{L}^1$ condition for integrability)

Given $E \in \mathcal{M}$ and $f \in \mathcal{M}e$ with

$$\int_E |f| dm < \infty, \tag{2.20}$$

we define

$$\int_E f dm = \int_E f_+ dm - \int_E f_- dm$$

Remark

We say that $f \in \mathcal{L}^1(E, dm)$ if (2.20) is satisfied.

Remark

This is the first example of an $\mathcal{L}^p$-space, for $p \geq 1$, and where

$$\mathcal{L}^p(E, dm) = \left\{ f \in \mathcal{M}e(E) : \int_E |f|^p dm < \infty \right\}$$

The case $p = 1$, with the case at hand, integrability, the case $p = 2$ leads to $\mathcal{L}^2$-theory, which is an Hilbert space (the only one among those Banach spaces).

Proposition 2.22 (Properties of $\mathcal{L}^1$ functions)

Given $f, g \in \mathcal{L}^1(E), E \in \mathcal{M}$, then

1. $cf \in \mathcal{L}^1(E)$ and $c \int_E f = \int_E cf$
2. $f + g \in \mathcal{L}^1(E)$ and $\int_E f + g = \int_E f + \int_E g$
3. if $f \leq g$ pointwise, then $\int_E f \leq \int_E g$

Proof. 1. and 2. are easy exercises, For 3., write $g - f \geq 0$, which imply $\int_E (g - f) \geq 0$ by monotonicity for non-negative measurable functions and $\int_E g - \int_E f$ by 2. ■

The following is a very important basic consequence

Corollary 2.23

Given $f \in \mathcal{L}^1(E), E \in \mathcal{M}$, then

$$\left| \int_E f \right| \leq \int_E |f|$$

Proof. Since $f \leq |f|$ and $-f \leq |f|$, which imply that

$$\max \left(\int_E f, -\int_E f \right) \leq \int_E |f|$$

■

Theorem 2.24 (Fatou's lemma)

Let $\{f_n\}_{n=1}^{\infty}$ a sequence of **non-negative** measurable functions. Define $f = \liminf_{n\to\infty} f_n \in \mathcal{M}e$; then

$$\int_E f dm \leq \liminf_{n\to\infty} \int_E f_n$$

that is

$$\int_E \left(\liminf_{n\to\infty} f_n \right) dm \leq \liminf_{n\to\infty} \left(\int_E f_n dm \right)$$

Proof. Let $g_k = \inf_{n\geq k} f_n$ and $a_k = \inf_{n\geq k}\int_E f_n$. Clearly, $0 \leq g_1 \leq g_2 \leq g_3 \leq \ldots$ and $0 \leq a_1 \leq a_2 \leq a_3 \ldots$. Apart from MCT, the basic inequality here is the following

$$a_k \geq \int_E g_k \tag{2.21}$$

since $g_k \leq f_n$ for $n \geq k$ and $f = \lim_{n\to\infty} g_k$ and $\lim_{n\to\infty} a_k = \liminf_{n\to\infty}\int_e f_n$, so

$$\begin{aligned}
\int_E f &= \int_E \lim_{k\to\infty} g_k \\
&= \lim_{k\to\infty}\int_E g_k && \text{(by MCT)} \\
&\leq \lim_{k\to\infty} a_k && \text{(by (2.21))} \\
&= \lim_{k\to\infty}\inf_{n\geq k}\int_E f_n \\
&= \liminf_{n\to\infty}\int_E f_n.
\end{aligned}$$

■

We give another application of the monotone convergence theorem (in this case to probability).

Lemma 2.25 (Borel-Cantelli)
Let $\{E_n\}_{n=1}^{\infty} \in \mathcal{M}$ with $\sum_{n=1}^{\infty} m(E_n) < \infty, E_n \subset E$. Denote

$$\{E_n; \text{i.o.}\} \coloneqq \bigcap_{k=1}^{\infty}\bigcup_{n\geq k} E_n \in \mathcal{M}e$$

Then $m(\{E_n; \text{i.o.}\}) = 0$.

Proof. Consider the characteristic functions of E_n,

$$\chi_{E_n}(x) = \begin{cases} 1 & \text{if } x \in E_n \\ 0 & \text{otherwise .} \end{cases}$$

where χ_{E_n} for $n = 1, \ldots$. Suppose $x \in \{E_n; \text{i.o.}\}$, this happens if and only if $\sum_{n=1}^{\infty}\chi_{E_n}(x) = \infty$. (exercise).

Let $g(x) = \sum_{n=1}^{\infty}\chi_{E_n}(x)$ and clearly $|g(x)| = g(x)$, since $g \geq 0$.

Claim

$g \in \mathcal{L}^1(E; dm)$. Then

$$\begin{aligned}\int_E |g| &= \int_E \left(\sum_{n=1}^{\infty} \chi_{E_n} \right) \\ &= \sum_{n=1}^{\infty} \int_E \chi_{E_n} \qquad \text{(by MCT)} \\ &= \sum_{n=1}^{\infty} m(E_n) \qquad \text{(finite by assumption)}\end{aligned}$$

Therefore $g \in \mathcal{L}^1(E; dm)$ implies $g(x) < \infty$ for a.e. $x \in E$, which implies

$$m(\{E_n; \text{i.o.}\}) = m(\{x; g(x) = \infty\}) = 0.$$

■

Theorem 2.26 (Dominated convergence (DCT))

Let $(f_n)_{n=1}^{\infty} \in \mathcal{M}_{\mathcal{E}}$. Assume

1. $f = \lim_{n\to\infty} f_n$ exists
2. $\exists g \in \mathcal{L}^1(E)$ with $|f_n| \leq g$ on E.

Then

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

Proof. Use Fatou's lemma applied in two ways: given $\{f_n\}_{n=1}^{\infty}$ and the dominating function $g \in \mathcal{L}^1(E)$, we form two **non-negative** sequences $\{g \pm f_n\}_{n=1}^{\infty}$. We apply Fatou to both sequences $\{g + f_n\}_{n=1}^{\infty}, \{g - f_n\}_{n=1}^{\infty}$; we have

$$\liminf_{n\to\infty} \int_E (g + f_n) \geq \int_E \liminf_{n\to\infty} (g + f_n)$$

and the left hand side can be written as

$$\int_E g + \liminf_{n\to\infty} \int_E f_n \geq \int_E g + \int_E \liminf_{n\to\infty} f_n$$

which implies

$$\liminf_{n\to\infty} \int_E f_n \geq + \int_E \liminf_{n\to\infty} f_n$$

Since by assumption $f = \lim_n f_n = \liminf_{n\to\infty} f_n$, this implies

$$\liminf_{n\to\infty} \int_E f_n \geq \int_E f \tag{2.22}$$

Note

Since $f \in \mathcal{L}^1$,

$$\int_E |f| = \int_E \liminf_{n\to\infty} |f_n| \leq \liminf_{n\to\infty} \int_E |f_n| \leq \int_E g < \infty$$

by Fatou, and since $g \in \mathcal{L}^1$. So $f \in \mathcal{L}^1$.

We now look at the other sequence $\{g - f_n\}_{n=1}^{\infty}$ and apply again Fatou

$$\liminf_{n\to\infty} \int_E (g - f_n) \geq \int_E \liminf_{n\to\infty} (g - f_n)$$

and since g does not depend on n, we can break this into

$$\int_E g + \liminf_{n\to\infty} \int_E (-f_n) \geq \int_E g + \int_E \liminf_{n\to\infty} (-f_n)$$

thus

$$-\liminf_{n\to\infty} \int_E (f_n) \geq -\int_E \limsup_{n\to\infty} f_n$$

This gives

$$\liminf_{n\to\infty} \left(-\int_E f_n \right) \geq -\int_E \limsup_{n\to\infty} (f_n)$$

therefore

$$\begin{aligned} -\limsup_{n\to\infty} \left(\int_E f_n \right) &\geq -\int_e \limsup_{n\to\infty} f_n \\ \Leftrightarrow \quad \limsup_{n\to\infty} \left(\int_E f_n \right) &\leq \int_E \limsup_{n\to\infty} f_n \end{aligned}$$

Since limits exists, $\limsup_{n\to\infty} f_n = f$, then

$$\limsup_{n\to\infty} \left(\int_E f_n \right) \leq \int_E f \tag{2.23}$$

Combining (2.22) and (2.23),

$$\limsup_{n\to\infty}\left(\int_E f_n\right) \leq \int_E f \leq \liminf_{n\to\infty}\left(\int_E f_n\right)$$

But $\liminf \leq \limsup$ **always**! This implies that

$$\limsup_{n\to\infty}\int_E f_n = \liminf_{n\to\infty}\int_E f_n = \lim_{n\to\infty}\int_E f_n$$

and moreover

$$\int_E = \limsup_{n\to\infty}\int_E f_n = \lim_{n\to\infty}\int_E = \liminf_{n\to\infty}\int_E f_n$$

■

2.5 Approximations of the identity

Given $f \in \mathcal{L}^p(\mathbb{R}^n)$ for $p \geq 1$, we want to approximate a rough function; the idea is to construct a sequence $\{\chi_\varepsilon\}$ nice smooth functions with the property that

$$\begin{aligned} f(x) &= \lim_{\varepsilon\to 0}\int_{\mathbb{R}^n}\chi_\varepsilon f \\ &= \int_{\mathbb{R}^n}\lim_{\varepsilon\to 0}\chi_\varepsilon f \end{aligned}$$

and we will show that $\lim_{\varepsilon\to 0}$

2.6 Riemann integral versus Lebesgue integral

Theorem 2.27

Let f be an integrable function over some $\Omega \subset \mathbb{R}^n$. Then Ω is measurable in the sense of Lebesgue and $f \in \mathcal{L}^1(\Omega)$ with

$$\int_\Omega f dm = \int_\Omega f(x)\mathrm{d}x$$

where the integral on the right is the Riemann integral.

Lebesgue integrable functions are also Riemann integrable functions if and only if the set of discontinuities should have Lebesgue measure zero. We demonstrate this with the Cantor staircase function.

Claim

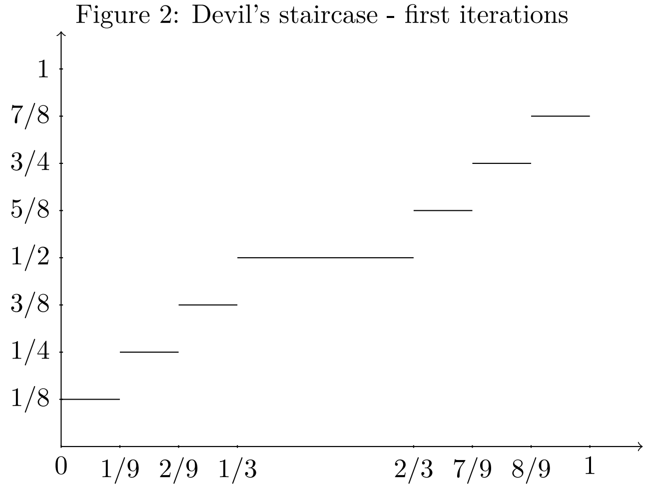
On $\mathcal{C}$, define ϕ by continuity. Our goal is to define $\int f(x)d\phi(x)$, the cantor set on the level of sets is self-similar.

$$\mathcal{C} = \frac{1}{3}(\mathcal{C}) + \frac{1}{3}(\mathcal{C} + 2)$$

where $x \mapsto \frac{x}{3}, x \mapsto \frac{x+2}{3}$ (an iterated function sequence) and we have

$$\phi\left(\frac{x}{3}\right) = \frac{1}{2}\phi(x)$$
$$\phi\left(\frac{x+2}{3}\right) = \frac{1}{2}\phi(x) + \frac{1}{2}$$

Figure 2: Devil's staircase - first iterations



We can compute the integral $\int_0^1 e^{ax} d\phi(x)$, the Lebesgue-Stieljes integral.

$$\begin{aligned} F(a) &= \int_0^1 e^{ax} d\phi(x) \\ &= \int_0^{\frac{1}{3}} e^{ax} d\phi(x) + \int_{\frac{2}{3}}^1 e^{ax} d\phi(x) \end{aligned}$$

let $x = \frac{y}{3}$ and in the second integral $x = \frac{y+2}{3}$

$$= \frac{1}{2}\left[\int_0^1 e^{a\frac{y_1}{3}}\, d\phi(y_1) + e^{\frac{2a}{3}} \int_0^1 e^{\frac{a y_2}{3}}\, d\phi(y_2)\right]$$

$$= \left[\frac{e^{\frac{2a}{3}}+1}{2}\right] F\left(\frac{a}{3}\right)$$

$$= e^{\frac{a}{3}} \cosh\left(\frac{a}{3}\right) F\left(\frac{a}{3}\right)$$

Now by induction

$$F(a) = \exp\left(a\left(\frac{1}{3}+\frac{1}{9}+\cdots+\frac{1}{3^k}\right)\right)\cosh\left(\frac{a}{3}\right)\times\cosh\left(\frac{a}{9}\right)\times\cdots\times\cosh\left(\frac{a}{3^k}\right)F\left(\frac{a}{3^k}\right)$$

and as $k \to \infty$, $F\left(\frac{a}{3^k}\right) \to F(0) = 1$, therefore

$$F(a) = e^{\frac{a}{2}} \prod_{k=1}^{\infty} \cosh$$

Now

$$\sin(\pi x) = \frac{e^{\pi i x} - e^{-\pi i x}}{2i}$$

and

$$I_2 = \frac{F(\pi i) - F(-\pi i)}{2i} \rightsquigarrow I_2 = \prod_{k=1}^{\infty} \cos\left(\frac{\pi}{3^k}\right)$$

One can show that $\mathcal{L}^1$ is not a Hilbert space, since the **parallelogram law** does not hold; there is no inner product for which

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 + \|\boldsymbol{x}-\boldsymbol{y}\|^2 = 2\left(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\right)$$

In $\mathcal{L}^p$, this becomes inequality, known as Hanner inequalities.[3] Recall that **Banach spaces** are complete normed vector spaces with $\mathcal{L}^p$ norm (over $\mathbb{R}, \mathbb{C}$ or anything reasonable). A space is an Hilbert space if and only if the distance comes from a norm, and in $p = 2$, this holds if an only if the parallelogram law holds. The case where $p = 2n$ is easier to deal with.

Recall from last semester that $\mathcal{L}^1$ is complete.

[3] For more, see Elliott H. Lieb and Michael Loss' book.

Definition 2.28 (Linear functional)
A **linear functional** is a map

$$A : \mathcal{X} \to \mathbb{R}$$
$$A(\alpha x + \beta y) = \alpha A(x) + \beta A(y)$$

for any two α, β constants and $x, y \in \mathcal{X}$, the Banach space. What is more interesting is to look at the continuous linear functions.
Definition 2.29 (Norm of linear functional)

$$A : \mathcal{X} \to \mathcal{Y}$$

where A is linear and where $\mathcal{X}, \mathcal{Y}$ are **normed linear spaces**. The norm of a linear functional A, sometimes called an operator norm, is

$$\|A\| := \sup_{\substack{x \in \mathcal{X} \\ \|x\| \leq 1}} \|Ax\|_{\mathcal{Y}} = \sup_{\|x\|_{\mathcal{X}} \neq 0} \frac{\|Ax\|_{\mathcal{Y}}}{\|x\|_{\mathcal{X}}}$$

and we say if $\|A\| < \infty$ that A is **bounded.**
Proposition 2.30
The following are equivalent:

(1) A is bounded

(2) A is continuous; if $\|x_n - x\|_{\mathcal{X}} \to 0$, then $\|Ax_n - Ax\|_{\mathcal{Y}} \to 0$

(3) A is continuous at a, a single point $x_0 \in \mathcal{X}$

Proof. The implication (1) $\Rightarrow$ (2):

$$\begin{aligned} \|Ax_1 - Ax_2\|_{\mathcal{Y}} &= \|A(x_1 - x_2)\|_{\mathcal{Y}} \\ &\leq \|A\|_{\text{op}} \|x_1 - x_2\|_{\mathcal{X}} \end{aligned}$$

provided $\|A\| < \infty$. (2) $\Rightarrow$ (3) is clear. For (3) $\Rightarrow$ (1), suppose A is continuous at $x_0 \in \mathcal{X}$; given $\varepsilon > 0, \exists \delta > 0$ such that

$$\|x - x_0\| < \delta \Rightarrow \|A(x - x_0)\| < \varepsilon$$

Let $\|h\| < \delta, h \in \mathcal{X}$. Then $\|x_0 + h - x_0\| < \delta$ implies

$$\|A(x_0 + h) - Ax_0\| = \|Ah\| < \varepsilon$$

■

Proposition 2.31 (Basic properties of $\mathcal{L}^1$)

1. $\mathcal{L}^1$ is a Banach space; it is a vector space that is complete in the $\|\cdot\|_{\mathcal{L}^1}$ norm. Given $\{f_k\}_{k=1}^{\infty} \in \mathcal{L}^1$ with $\|f_k - f\|_{\mathcal{L}^1} \to 0$ as $k \to \infty$ implies $f \in \mathcal{L}^1$.

2. Subsequential compactness in $\mathcal{L}^{\infty}$ *i.e.* suppose $f_n \to f$ in $\mathcal{L}^1$, (*i.e.* $\lim_{n\to\infty} \|f_n - f\|_{\mathcal{L}^1} = 0$)[4] Then , there exists a subsequence $\{f_{n_k}\}_{k=1}^{\infty}$ with the property that $f_{n_k}(x) \to f(x)$ for a.e. $x \in \mathbb{R}^n$.

3. Density: step functions, simple functions, continuous functions of compact support (denoted $\mathcal{C}_0^0$) are all dense in $\mathcal{L}^1$

2.7 Fubini theorem

Take $\mathbb{R}^d = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, the product space for $d_1, d_2 \geq 1$ and consider

$$f(x, y) \in \mathcal{M}e(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2})$$

We form a slice function

$$\begin{aligned} f^y(x) &:= f(x, y); \qquad x \in \mathbb{R}^{d_1} \\ f^x(y) &:= f(x, y); \qquad y \in \mathbb{R}^{d_2} \end{aligned}$$

Given a set $E \subset \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, we form the corresponding slices

$$\begin{aligned} E^y &= \{x \in \mathbb{R}^{d_1} : (x, y) \in E\} \\ E^x &= \{y \in \mathbb{R}^{d_2}; (x, y) \in E\} \end{aligned}$$

Theorem 2.32 (Fubini's theorem)

Suppose $f(x, y) \in \mathcal{L}^1(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2})$. Then, for almost every $y \in \mathbb{R}^{d_2}$,

1. $f^y \in \mathcal{L}^1(\mathbb{R}^{d_1})$

2. $\int_{\mathbb{R}^{d_1}} f^y(x) \mathrm{d}x \in \mathcal{L}^1(\mathbb{R}^{d_2})$ and moreover

$$\int_{\mathbb{R}^{d_2}} \left(\int_{\mathbb{R}^{d_1}} f(x, y) \mathrm{d}x \right) \mathrm{d}y = \int_{\mathbb{R}^d} f$$

[4] This does not imply pointwise, or even almost everywhere pointwise convergence.

Note

There is a symmetry in x and y, that is $f^x \in \mathbb{R}^{d_2}$ and $\int_{\mathbb{R}^{d_2}} f^x(y)\mathrm{d}y \in \mathcal{L}^1(\mathbb{R}^{d_1})$ with

$$\int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f(x,y)\mathrm{d}y \right) \mathrm{d}x = \int_{\mathbb{R}^d} f$$

Proof. [Schematic] The basic idea is to prove that for a given E measurable, we have for a.e. y, E^y is measurable and same for x.

Here is the rough outline: let $\mathcal{F} = \{f \in \mathcal{L}^1(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}); f(x,y) \text{ satisfies Fubini}\}$

1. Given $\{f_k\}_{k=1}^n \in \mathcal{F}$, linear combinations of $\{f_k\}_{k=1}^n$ are $\int f_k^y \mathrm{d}x$ in $\mathcal{F}$ (by linearity of integrals).
2. Suppose $\{f_k\} \in \mathcal{Me} \subset \mathcal{F}$ and that $f_k \nearrow f$ (or $f_k \searrow f$). Then, by MCT (exercise), one gets that $f \in \mathcal{F}$.
3. Now, we build up a progression of functions $f \in \mathcal{F}$.
   (1) Suppose $Q \subset \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ open cube, $\chi_Q \in \mathcal{L}^1(\mathbb{R}^d)$ and $\chi_Q \in \mathcal{F}$ (volume of a cube).
   (2) The boundary of the cubes ∂Q has connected components, $\chi_{\partial Q} \in \mathcal{F}$.
   (3) Let $E \subset \mathbb{R}^d$ open with $m(E) < \infty$. Then, we know that $E = \bigcup_{j=1}^{\infty} Q_j$ almost disjoint union. We build up χ_E using $f_k = \sum_{j=1}^{k} \chi_{Q_j}$, increasing in k so that $f_k \nearrow \chi_E$ monotonically, and since $f_k \in \mathcal{F}$, therefore by monotone convergence theorem $\chi_E \in \mathcal{F}$.
   (4) Now, we go to G_δ sets: $E \in G_\delta$ (countable intersection of open sets), $E = \bigcap_{k=1}^{\infty} O_k$ for O_k open with $m(O_k) < \infty$. We approximate χ_E by $\chi_{\cap_{k=1}^n O_k}$ as $n \to \infty$, $\chi_{\cap_{k=1}^n O_k} \searrow \chi_E$ as $n \to \infty$, $\chi_E \in \mathcal{F}$
   (5) Show that if $E \in \mathcal{M}$ with $m(E) = 0$, then $E \in \mathcal{F}$ (exercise).
   (6) Since $\mathcal{M} = G_\delta \cup \{\text{measure zero}\}$, we have that $\chi_E \in \mathcal{F}$ for any $E \in \mathcal{M}$ with $m(E) < \infty$.
   (7) Given $f \in \mathcal{Me} \cap \mathcal{L}^1(\mathbb{R}^n)$, we can find simple functions $\{\phi_k\}_{k=1}^{\infty}$ with

   $$\phi_k = \sum_{j=1}^{k} a_j \chi_{E_j},$$

   $E_j \in \mathcal{M}, m(E_j) < \infty$ with $\phi_k \nearrow f$. Since $\phi_k \in \mathcal{F}$ for all k, by monotonicity (by MCT), we get $f \in \mathcal{F}$.

■

The problem of this is that it is sometimes hard to verify that $f \in \mathcal{L}^1$.

Theorem 2.33 (Fubini-Tonelli)

Suppose $f(x,y) \in \mathcal{M}$ and $f(x,y) \geq 0$. Then, for almost every $y \in \mathbb{R}^{d_1}$,

1. $f^y \in \mathcal{M}e(\mathbb{R}^{d_1})$
2. $\int_{\mathbb{R}^{d_1}} f^y(x)\mathrm{d}x \in \mathcal{M}e(\mathbb{R}^{d_2})$
3. $$\int_{\mathbb{R}^{d_2}} \left(\int_{\mathbb{R}^{d_1}} f(x,y)\mathrm{d}x \right) \mathrm{d}y = \int_{\mathbb{R}^d} f$$

   and the same is true for f^x replacing f^y. So, in particular,

   $$\int_{\mathbb{R}^{d_2}} \left(\int_{\mathbb{R}^{d_1}} f(x,y)\mathrm{d}x \right) \mathrm{d}y = \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f(x,y)\mathrm{d}y \right) \mathrm{d}x = \int_{\mathbb{R}^d} f$$

   unconditionally, but could have $\infty = \infty$.

Note

We do not require that $f \in \mathcal{L}^1$ here!

Note

It thus suffices to take absolute values and check for iterated integrals for the absolute value of f. The practical implication is the following; given $f \in \mathcal{M}e(\mathbb{R}^d)$, we consider $|f| \in \mathcal{M}e(\mathbb{R}^d)$ and apply Tonelli (theorem 2.33) to $|f|$. If the iterated integrals $\int \left(\int |f|\mathrm{d}x\right) \mathrm{d}y < \infty$, then $f \in \mathcal{L}^1$ and we can apply Fubini's theorem.

Proof. The idea is to construct monotone sequences $\{f_k(x,y)\}_{k=1}^{\infty}$ converging to $f(x,y)$ and use MCT. For instance, one can put

$$f_k(x,y) = \begin{cases} f(x,y) & \text{if } |x,y| < k \text{ and } |f(x,y)| < k \\ 0 & \text{otherwise} \end{cases}$$

and clearly, $f_k \in \mathcal{L}^1(\mathbb{R}^{d_1} \times \mathbb{R}_{d_2})$ (bounded function of compact support), therefore the exists $E_k \subset \mathbb{R}^{d_2}$ with $m(E_k) = 0$ such that $f_k^y(x) \in \mathcal{M}e$ for $y \in E_k^{\complement}$. Now, we let $E = \bigcup_{k=1}^{\infty} E_k$. Clearly, $m(E) = 0$; this implies that $f^y(x) \in \mathcal{M}e$ for all $y \in E^{\complement}$ (here we use that $f_k^y \nearrow f^y$). Since $f_k^y \nearrow f^y$ as $k \to \infty$(non-negative), by the MCT, if $y \notin E$,

$$\int_{\mathbb{R}^{d_1}} f_k(x,y)\mathrm{d}x \nearrow_{k\to\infty} \int_{\mathbb{R}^{d_1}} f(x,y)\mathrm{d}x \tag{2.24}$$

Again, by Fubini, for any $k \geq 1$,

$$\int_{\mathbb{R}^{d_1}} f(x,y)\mathrm{d}x \in \mathcal{M}e(\mathbb{R}^{d_2})$$

for all $y \in E^{\complement}$ and all $y \in E^{\complement}$. Apply Fubini again to the $f_k's$

$$\int_{\mathbb{R}^{d_2}} \left(\int_{\mathbb{R}^{d_1}} f_k(x,y)\mathrm{d}x \right) \mathrm{d}y = \int_{\mathbb{R}^d} f_k \tag{2.25}$$

We apply MCT again to (2.25) and combine with (2.24) to get

$$\int_{\mathbb{R}^{d_2}} \left(\int_{\mathbb{R}^{d_1}} f(x,y)\mathrm{d}x \right) \mathrm{d}y = \int_{\mathbb{R}^d} f$$

■

An immediate corollary of Fubini-Tonnelli,

Corollary 2.34

For any $E \in \mathcal{M}(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2})$, we have for a.e. $y \in \mathbb{R}^{d_2}$

$$E^y = \{x \in \mathbb{R}^{d_1} | (x,y) \in E\} \in \mathcal{M}(\mathbb{R}^{d_1})$$

Moreover, $m(E^y) \in \mathcal{M}e(\mathbb{R}^{d_2})$ and $m(E) = \int_{\mathbb{R}^{d_2}} m(E^y)\mathrm{d}y$.

Proof. Just apply Fubini-Tonnelli with $f(x,y) = \chi_E(x,y)$. ■

Section 3
Hilbert spaces

Informally, a Hilbert space is the infinite dimensional generalizations of finite-dimensional vector spaces. $\mathcal{H}$ has an inner product $\langle\cdot,\cdot\rangle$ generalizing the usual inner product on $\mathbb{R}^n$. Many analogies between $\mathcal{H}$ and what you learned in linear algebra hold.[5]

Example 3.1

1. $$\mathcal{L}^2(\mathbb{R}^n) = \left\{ f \in \mathcal{M}e(\mathbb{R}^n); \int_{\mathbb{R}^n} |f|^2 \mathrm{d}x < \infty \right\}$$

2. $$\ell^2 = \left\{ (a_k)_{k=1}^{\infty}; a_k \in \mathbb{C}, \sum_{k=1}^{\infty} |a_k|^2 < \infty \right\}.$$

   There is a direct analog between ℓ^2 and Fourier series,

   $$\mathcal{L}^2([-\pi,\pi]) = \{f | \int_{-]pi}^{\pi} |f|^2 < \infty\}$$

3. Sobolev spaces, $H^s, s \in \mathbb{R}$, with

   $$H^s = \{f : \partial^k f \in \mathcal{L}^2 \ \forall\ k \leq s\}$$

   that is if $X \in \mathcal{C}_0^\infty$, we have the distributional derivative definition $\langle \partial^k f, X\rangle_{\mathcal{L}^2} := \langle f, (-1)^k \partial^k X\rangle$.

4. Hardy spaces

We start by looking at $\mathcal{L}^2(\mathbb{R}^n)$. There is a **norm**

$$\|f\|_{\mathcal{L}^2} := \left(\int_{\mathbb{R}^n} |f|^2 \mathrm{d}x\right)^{\frac{1}{2}}.$$

However, there is also an inner-product,

$$\langle f, g\rangle_{\mathcal{L}^2} := \int_{\mathbb{R}^n} f(x)\overline{g(x)}\mathrm{d}x$$

[5]Recall that a Banach space is a complete normed vector space. A Hilbert space is a Banach space with the norm induced by an inner product, which allows one to measure angles and not just measure and distance. $\mathcal{L}^2$ has a unique structure that distinguish it from $\mathcal{L}^p$ for $p \neq 2$..

Note

1. First, $\langle f, f\rangle_{\mathcal{L}^2} = \int_{\mathbb{R}} |f(x)|^2 \mathrm{d}x = \|f\|^2_{\mathcal{L}^2}$
2. Clearly, $\langle \cdot, \cdot \rangle : \mathcal{L}^2 \times \mathcal{L}^2 \to \mathbb{R}$ is bilinear because the integral is.

Moreover, $\mathcal{L}^2$ has the following properties,

Proposition 3.1

1. $\mathcal{L}^2$ is a vector space
2. $f(x)\overline{g(x)} \in \mathcal{L}^1(\mathbb{R}^n)$ if $f, g \in \mathcal{L}^2(\mathbb{R}^n)$.We also have the Cauchy-Schwarz inequality,

$$|\langle f, g\rangle| \leq \|f\|_{\mathcal{L}^2}\|g\|_{\mathcal{L}^2}$$

Proof.

1. If $f \in \mathcal{L}^2$ and $\alpha \in \mathbb{C}$, then $\alpha f \in \mathcal{L}^2$. So enough to show that for $f, g \in \mathcal{L}^2$, then $f + g \in \mathcal{L}^2$. We show that

$$\|f + g\|_{\mathcal{L}^2} \leq \|f\|_{\mathcal{L}^2} + \|g\|_{\mathcal{L}^2},$$

the traingle inequality. To see this, we note that

$$|f + g|^2 \leq 4(|f|^2 + |g|^2) \tag{3.26}$$

since $2|fg| \leq |f|^2 + |g|^2$. So from (3.26), we have

$$\int_{\mathbb{R}^n} |f + g|^2 \mathrm{d}x \leq 4\left(\int_{\mathbb{R}^n} |f|^2 + \int_{\mathbb{R}^n} |g|^2\right) < \infty$$

2. We have

$$|fg| \leq \frac{1}{2}(|f|^2 + |g|^2)$$
$$\Rightarrow \quad \int_{\mathbb{R}^n} |f\overline{g}| \leq \frac{1}{2}\left(\int_{\mathbb{R}^n} |f|^2 + \int_{\mathbb{R}^n} |g|^2\right)$$

implies that $f(x)\overline{g(x)} \in \mathcal{L}^1(\mathbb{R}^n)$. For Cauchy-Schwarz, when either $f = 0$ or $g = 0$, this is obvious so assume $f \neq 0$ and $g \neq 0$ a.e. Now, consider $F = f/\|f\|_{\mathcal{L}^2}$ and $G = g/\|g\|_{\mathcal{L}^2}$. The Cauchy-Schwarz for F and G is

$$|\langle F, G\rangle| \leq \|F\|_{\mathcal{L}^2}\|G\|_{\mathcal{L}^2} = 1$$

Now using the same quadratic formula,

$$|F\overline{G}| \leq \frac{1}{2}\left(|F|^2 + |G|^2\right) \tag{3.27}$$

and using (3.27),

$$\begin{aligned} |\langle F, G\rangle| &\leq \int_{\mathbb{R}^n} |F\overline{G}|\mathrm{d}x \\ &\leq \frac{1}{2}\left(\int_{\mathbb{R}^n} |F|^2\mathrm{d}x + \int_{\mathbb{R}^n} |G|^2\mathrm{d}x\right) \end{aligned}$$

and so

$$\left|\left\langle \frac{f}{\|f\|}, \frac{g}{\|g\|} \right\rangle\right| \leq 1$$

if and only if

$$|\langle f, g\rangle| \leq \|f\|\|g\|$$

■

Last time, we proved some basic properties regarding $\mathcal{L}^2$ and the inner product $\langle f, g, \rangle = \int_{\mathbb{R}^n} f\overline{g}\mathrm{d}x$ last time, for $f, g \in \mathcal{L}^2(\mathbb{R}^n)$. There are two important properties that we have to establish: **completeness** and **separability**

Theorem 3.2 (Compleness of $\mathcal{L}^2$)
$\mathcal{L}^2(\mathbb{R}^n)$ is complete, namely Cauchy sequences converge in $\mathcal{L}^2(\mathbb{R}^n)$.

Proof. Parallels the proof in $\mathcal{L}^1$: the key point is that the triangle inequality holds. As in the $\mathcal{L}^1$ case, given a Cauchy sequence $\{f_n\}_{n=1}^{\infty}$, we choose a subsequence $\{f_{n_k}\}$ that converges very quickly so that $\|f_{n_{k+1}} - f_{n_k}\| \leq 2^{-k}$ for $k = 1, 2, \ldots$ where $\|\cdot\|$ denotes the $\mathcal{L}^2$ norm. Set

$$f(x) = f_{n_1}(x) + \sum_{k=1}^{\infty}\left(f_{n_{k+1}}(x) - f_{n_k}(x)\right) \tag{3.28}$$

$$g(x) = |f_{n_1}(x)| + \sum_{k=1}^{\infty}\left|f_{n_{k+1}}(x) - f_{n_k}(x)\right|$$

Similarly, define the corresponding partial sums

$$S_{N(f)}(x) = f_{n_1}(x) + \sum_{k=1}^{N} \left(f_{n_{k+1}}(x) - f_{n_k}(x)\right) \tag{3.29}$$

$$S_{N(g)}(x) = |f_{n_1}(x)| + \sum_{k=1}^{N} \left|f_{n_{k+1}}(x) - f_{n_k}(x)\right|$$

We have

$$\begin{aligned} \|S_{N(g)}\|_{\mathcal{L}^2} &\leq \|f_{n_1}\| + \sum_{k=1}^{N} \left\|f_{n_{k+1}}(x) - f_{n_k}(x)\right\| \\ &\leq \|f_{n_1}\| + \sum_{k=1}^{N} 2^{-k} \\ &\leq \|f_{n_1}\| + 1, \end{aligned}$$

which is the dominating function and since by construction (usinga telescoping sum), then $\lim_{N\to\infty} S_{N(g)} = g$ **pointwise** (**monotone**, non-negative convergence), then by the monotone convergence theorem (or DCT), we get $\|g\|_{\mathcal{L}^2} < \infty$. ■

Note

Since $|f(x)| \leq g$, in particular the fact that $\|g\|_{\mathcal{L}^2} < \infty$ implies that $f \in \mathcal{L}^2(\mathbb{R}^n)$, which in terms implies that the series defining $f(x)$ in (3.28) is finite a.e. m (converges almost everywhere). This means that $f(x)$ defined in (3.28) is indeed a good candidate for limit in $\mathcal{L}^2$.

Claim

$\lim_{k\to\infty} \|f_{n_k} - f\|_{\mathcal{L}^2} = 0$

Proof. The key point is that $f_{n_k}(x) = S_{k-1(f)}(x)$. Since $S_{k-1(f)}(x) \xrightarrow{k\to\infty} f(x)$ for a.e. $x \in \mathbb{R}^n$. This implies that $f_{n_k}(x) \xrightarrow{k\to\infty} f(x)$ for a.e. $x \in \mathbb{R}^n$. We apply DCT to show that

$$\lim_{k\to\infty} \|f_{n_k} - f\|_{\mathcal{L}^2} = 0$$

Now

$$\begin{aligned} \|f_{n_k} - f\|_{\mathcal{L}^2} &= \|S_{k-1(f)} - f\|_{\mathcal{L}^2} \\ &\leq \|S_{k-1(f)}\|_{\mathcal{L}^2} + \|f\|_{\mathcal{L}^2} \\ &\leq \|g\|_{\mathcal{L}^2} + \|g\|_{\mathcal{L}^2} = 2\|g\|_{\mathcal{L}^2} < \infty \end{aligned} \tag{3.30}$$

Since $\lim_{k\to\infty} f_{n_k}(x) = f(x)$ for a.e. $x \in \mathbb{R}^n$, from (3.30), we get by the DCT,

$$\lim_{k\to\infty} \|f_{n_k} - f\|_{\mathcal{L}^2} = \left\| \lim_{k\to\infty} (f_{n_k} - f) \right\|_{\mathcal{L}^2} = 0$$

So far, we have completeness for such subsequences $\{f_{n_k}\}_{k=1}^{\infty}$. This is **not** a restriction. Choose $\varepsilon > 0$ arbitrary,

$$\|f_n - f\|_{\mathcal{L}^2} \leq \|f_n - f_{n_k}\|_{\mathcal{L}^2} + \|f_{n_k} - f\|_{\mathcal{L}^2}$$

and we have already show that for $n_k \geq M(\varepsilon)$ large enough, $\|f_{n_k} - f\|_{\mathcal{L}^2} < \frac{\varepsilon}{2}$. On the other hand, by Cauchy condition,

$$\|f_n - f_{n_k}\|_{\mathcal{L}^2} \leq \frac{\varepsilon}{2} \quad \text{if } n, n_k > N$$

So chose $\widetilde{N} = \max(N, M)$ and pick $n > \widetilde{N}$. ■

Another important property is **separability**.

Theorem 3.3

$\mathcal{L}^2(\mathbb{R}^n)$ is separable (*i.e.* there is a countable collection of $\mathcal{L}^2$ functions whose linear combinations are dense in $\mathcal{L}^2(\mathbb{R}^n)$.)

Proof. We have to use $\mathbb{Q}^n$. Consider the functions

$$\{r\chi_R(x)\}_{\substack{R\in\mathbb{Q}^n \\ r\in\mathbb{C}}}.$$

Here, χ_R is characteristic function of rectangle R with rational coordinates [6] The problem is that $\mathcal{L}^2(\mathbb{R}^n) \not\subseteq \mathcal{L}^1(\mathbb{R}^n)$.

Step 1: Approximate $f \in \mathcal{L}^2$ by an $\mathcal{L}^1$ function on a large ball. Take

$$g_n(x) = \begin{cases} f(x); & |x| < n \text{ and } |f(x) < n \\ 0 & \text{otherwise} \end{cases}$$

The first point to notice is $g_n(x) \xrightarrow{n\to\infty} f(x)$ for a.e $x \in \mathbb{R}^n$ (exercise). The second point is that for each $n \geq 1, g_n \in \mathcal{L}^1(\mathbb{R}^n)$.

[6] In $\mathcal{L}^1$, step functions are dense and it suffices to refine to rational coordinates. This is easy, but not so trivial in $\mathcal{L}^2$.

Note

We have

$$|g_n - f|^2 \leq (2|f|)^2 = 4|f|^2$$
$$\Rightarrow \|g_n - f\|_{\mathcal{L}^2} \leq \|f\|_{\mathcal{L}^2} < \infty.$$

Since $g_n(x) \to f(x)$ for a.e. $x \in \mathbb{R}^n$. Then by DCT,

$$\lim_{n\to\infty} \|g_n - f\|_{\mathcal{L}^2} = \left\| \lim_{n\to\infty} (g_n - f) \right\|_{\mathcal{L}^2} = 0.$$

So, given $\varepsilon > 0$, we can find $N = N(\varepsilon) > 0$ so that $\|f - g_N\|_{\mathcal{L}^2} < \frac{\varepsilon}{2}$. Let $g = g_N \in \mathcal{L}^1(\mathbb{R}^n)$; we can find a step function φ with $|\varphi| \leq N$ and

$$\int_{\mathbb{R}^n} |g - \varphi| \mathrm{d}x = \|g - \varphi\|_{\mathcal{L}^1(\mathbb{R}^n)} \leq \frac{\varepsilon^2}{16N}$$

By replacing φ with step function ψ with rational coordinates so that $|\psi| \leq N$ and

$$\|\psi - \varphi\|_{\mathcal{L}^1} \leq \frac{\varepsilon^2}{8N}.$$

■

We want to estimate

$$\begin{aligned}
\|g - \psi\|_{\mathcal{L}^2(\mathbb{R}^n)} &= \int_{\mathbb{R}^n} |g - \psi||g - \psi| \mathrm{d}x \\
&\leq \sup_{x \in \mathbb{R}^n} |g(x) - \psi(x)| \|g - \psi\|_{\mathcal{L}^1(\mathbb{R}^n)} \\
&\leq 2N \frac{\varepsilon^2}{8N}
\end{aligned}$$

which implies that in $\mathcal{L}^2$,

$$\|f - \psi\|_{\mathcal{L}^2} \leq \|f - g\|_{\mathcal{L}^2} + \|g - \psi\|_{\mathcal{L}^2} \leq \varepsilon + \frac{\varepsilon^2}{8N}$$

using the triangle inequality.

3.1 Hilbert spaces

We now discuss more general Hilbert spaces.

Definition 3.4 (Hilbert space)

A Hilbert space $\mathcal{H}$ is a vector space over $\mathbb{C}$ (or $\mathbb{R}$) with some properties.

1. There is an inner product $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ such that

   (a) $f \mapsto \langle f, g \rangle$ is linear for fixed $g \in \mathcal{H}$.

   (b) $\langle f, g \rangle = \overline{\langle g, f \rangle}$

   (c) $\langle f, f \rangle = \|f\|^2 \geq 0$.

   (d) $\|f\| = 0 \Rightarrow f = 0$.

2. $\mathcal{H}$ is complete in the metric $d(f, g) = \|f - g\|$

3. $\mathcal{H}$ is separable (it has a countable dense subset).

Remark

The triangle inequality $\|f + g\| \leq \|f\| + \|g\|$ and Cauchy-Schwarz $\langle f, g \rangle \leq \|f\|\|g\|$ are easy consequences of (a) - (d) using the same argument as for $\mathcal{L}^2(\mathbb{R}^n)$

Some examples

Example 3.2

1. If $E \in \mathcal{M}(\mathbb{R}^n)$ measurable with $m(E) > 0$, then $\mathcal{L}^2(E, \mathrm{d}x) = \mathcal{H}$ is a Hilbert space with $\mathrm{d}x$ the Lebesgue measure and $\langle f, g \rangle = \int_E f(x)\overline{g(x)}\mathrm{d}x$ and

$$\langle f, f \rangle = \|f\|^2_{\mathcal{L}^2} = \int_E |f(x)|^2 \mathrm{d}x.$$

   The case $E = [-\pi, \pi] \subset \mathbb{R}$ is of special significance (Fourier series).

2. $\mathbb{C}^N$ or $\mathbb{R}^N$ with the usual inner product (finite dimensional vector spaces). In this case, finite-dimensional basis implies separability.

3. $\ell^2(\mathbb{Z})$, defined by

$$\ell^2(\mathbb{Z}) := \left\{ (a_k)_{k=-\infty}^{\infty}; a_k \in \mathbb{C} \text{ with } \sum_{k=-\infty}^{\infty} |a_k|^2 \right\}$$

   and

$$\langle a, b \rangle = \sum_{k=-\infty}^{\infty} a_k \overline{b_k}$$

   for $(a_k)_{k=-\infty}^{\infty}, (b_k)_{k=-\infty}^{\infty}$. $\ell^2(\mathbb{Z})$ is also a Hilbert space (exercise). One can just as well do this for $\ell^2(\mathbb{N})$.

3.2 Orthogonality

Definition 3.5 (Orthogonality and orthonormality)

1. Given $f, g \in \mathcal{H}$ we can say that f is orthogonal to g if $\langle f, g, \rangle = 0$ (denoted $f \perp g$).

2. A sequence $\{e_k\}_{k=1}^{\infty} \subset \mathcal{H}$ is **orthonormal** provided

$$\langle e_k, e_l \rangle = \delta_l^k = \begin{cases} 1 & \text{if } k = l \\ 0 & \text{otherwise} \end{cases}$$

Proposition 3.6

If $\{e_k\}_{k=1}^{N}$ is any **finite** collection of orthonormal vectors with

$$f = \sum_{k=1}^{N} a_k e_k; \quad a_k \in \mathbb{C}$$

Then $\|f\|^2 = \sum_{k=1}^{N} |a_k|^2$ and the proof relies on Pythagoras theorem

We want to generalize this "basis" to infinite sets of e_k's.

Definition 3.7 (Hilbert basis)

An orthonormal subset $\{e_k\}_{k=1}^{\infty} \subset \mathcal{H}$ is called an **orthonormal basis** (or **Hilbert basis**) if finite liear combinations of e_k's are dense in $\mathcal{H}$.

Theorem 3.8

Every Hilbert space $\mathcal{H}$ has a Hilbert basis.

Proof. Given a countable dense subset $\{g_k\}_{k=1}^{\infty}$ we create independence by throwing away dependent vectors $\{h_l\}_{l=1}^{\infty}$ and using Gram-Schmidt, you gets $\{e_m\}_{m=1}^{\infty}$. ■

We want to understand in more detail how Hilbert basis mimic orthonormal basis in finite dimension.

Remark

Given $f \in \mathcal{H}$, $\{e_k\}_{k=1}^{\infty} \subset \mathcal{H}$, we often write "$f = \sum_{k=1}^{\infty} a_k e_k$", this does not mean pointwise equality. This means

$$\lim_{N \to \infty} \left\| f - \sum_{k=1}^{N} a_k e_k \right\| = 0$$

Theorem 3.9

The following are equivalent: given an orthonormal set $\{e_k\}_{k=1}^{\infty}$

1. $\{e_k\}_{k=1}^{\infty}$ is a Hilbert basis
2. If $f \in \mathcal{H}$ and $\langle f, e_k \rangle = 0$ for all $k = 1, 2, \ldots$, then $f = 0$.
3. If $f \in \mathcal{H}$ and $S_N(f) = \sum_{k=1}^{N} \langle f, e_k \rangle e_k$ for S_N the partial Fourier series. Then

$$\lim_{N \to \infty} \|f - S_N(f)\| = 0$$

4. If $a_k = \langle f, e_k \rangle$ for $k = 1, 2, \ldots$, then

$$\|f\| = \sum_{k=1}^{\infty} |a_k|^2,$$

the **Parseval identity**.[7]

Proof.

(1) $\Rightarrow$ (2): There is a subset $\{g_n\}_{n=1}^{\infty} \subset \mathcal{H}$ that is dense in $\mathcal{H}$ and g_n's are finite linear combinations of e_k's. So $g_n = \sum_{k=1}^{n} c_k e_k$, where $c_k \in \mathbb{C}$. In other words, given f, can find such g_n that approximates f to arbitrary accuracy with $\|g_n - f\| \to 0$ as $n \to \infty$.

Note

Since $\langle f, e_k \rangle = 0$ with all e_k's, then this implies $\langle f, g_n \rangle = 0$ for all n by linearity.

$$\begin{aligned} \|f\|^2 &= \langle f, f - g_n \rangle \quad \forall\, n \\ &\leq \|f\| \|f - g_n\| \end{aligned}$$

by Cauchy-Schwarz with $\|f - g_n\| \to 0$ as $n \to \infty$. Let $n \to \infty$, then this implies $\|f\|^2 = 0$ which implies that $\|f\| = 0$ such that $f = 0$.

(2) $\Rightarrow$ (3): Let $S_N(f) = \sum_{k=1}^{N} \langle f, e_k \rangle e_k$, the generalized N^{th} Fourier partial sum. We can write $f = S_N(f) + (f - S_N(f))$, which is an orthogonal decomposition. Indeed, if we look at the inner product

$$\langle f - S_N(f), S_N(f) \rangle = \left\langle f, \sum_{k=1}^{N} \langle f, e_k \rangle e_k \right\rangle - \sum_{k=1}^{N} \langle f_j, e_k \rangle^2 = 0$$

and Pythagoras theorem implies

$$\|f\|^2 = \|f - S_N(f)\|^2 + \|S_N(f)\|^2.$$

We want to show that under assumption (2), the second term goes to zero.

$$= \|f - S_N(f)\|^2 + \sum_{k=1}^{N} |a_k|^2 \tag{3.31}$$

where $a_k = \langle f, e_k \rangle$.

[7] Note that a collection of independent orthonormal vectors are not necessarily a Hilbert basis and so this property needs not hold

Remark

Equation (3.31) and taking $N \to \infty$ implies that

$$\sum_{k=1}^{\infty} |a_k|^2 \leq \|f\|^2$$

the **Bessel inequality**, which is always true indepndent of whether $\{e_k\}_{k=1}^{\infty}$ is a Hilbert basis.

We first show that $S_N(f)$ converge in $\|\cdot\|$. We do this by proving that $\{S_N(f)\}_{N=1}^{\infty}$ is Cauchy in $\|\cdot\|$. Assume $N > M$,

$$\begin{aligned} \|S_N(f) - S_M(f)\| &= \left\| \sum_{k=M+1}^{N} a_k e_k \right\| \\ &\leq \sum_{k=N+!}^{M} |a_k|^2 \end{aligned}$$

by Cauchy-Schwarz inequality. Since $f \in \mathcal{H}$, and by Bessel inequality, we have

$$\sum_{k=1}^{\infty} |a_k|^2 < \infty$$

So as $M, N \to \infty$, this implies

$$\sum_{k=M+1}^{N} |a_k|^2 \to 0$$

therefore $\{S_N(f)\}_{N=1}^{\infty}$ is Cauchy. Consequently, there exists $g \in \mathcal{H}$ such that $\|S_N(f) - g\| \to 0$ as $N \to \infty$. Fix $j \geq 1$; then for large N, $\langle f - S_N(f), e_j \rangle = 0$. Since $f \to g$ in $\mathcal{H}$, thus

$$\langle f - g, e_j \rangle = \langle f - S_N(f), e_j \rangle + \langle S_N(f) - g, e_j \rangle$$

Thus for $N >> j$ large, so $\langle f - S_N(f), e_j \rangle$ is zero and by Cauchy-Schwarz,

$$|\langle S_N(f) - g, e_j \rangle| \leq \|S_N(f) - g\| \to 0$$

as $N \to \infty$. This implies that $\langle f - g, e_j \rangle$ for all $j = 1, 2, \ldots$. Using the assumption (2), $f = g$ (as vectors in $\mathcal{H}$).

(3) $\Rightarrow$ (4) To prove Parseval identity using (3), write $f = (f - S_N(f)) + S_N(f)$. Since (3.31) is an orthogonal decomposition,

$$\|f\|^2 = \|f - S_N(f)\|^2 + \|S_N(f)\|^2.$$

By (3), $\|f - S_N(f)\| \to 0$ as $N \to \infty$. Thus,

$$\|f\|^2 = \sum_{k=1}^{\infty} |a_k|^2$$

(4) $\Rightarrow$ (1): Simply use that assuming Parseval, $\|f - S_N(f)\|^2 \to 0$ as $N \to \infty$. Since $\sum_{k=1}^{\infty} |a_k|^2 < \infty$. But $S_N(f) = \sum_{k=1}^{N} a_k e_k$, finite linear combination of basis elements. ∎

3.3 Fourier series

Consider the Hilbert space $\mathcal{H} = \mathcal{L}^2([-\pi, \pi])$, with $dm = \mathrm{d}x/2\pi$. The inner product for $f, g \in \mathcal{L}^2([-\pi, \pi])$ is

$$\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\overline{g(x)}\mathrm{d}x.$$

There is a **distinguished orthonormal** set

$$\{e^{inx}\}_{n \in \mathbb{Z}} = \{e^{inx}\}_{n=-\infty}^{\infty}$$

where the Euler identity

$$e^{inx} = \cos(nx) + i \sin(nx).$$

Also, it is convenient (but not necessary) to assume that $f(-\pi) = f(\pi)$. A crucial fact (which is not obvious) is the fact that $\{e^{inx}\}_{n \in \mathbb{Z}}$ is a Hilbert basis for $\mathcal{L}^2([-\pi, \pi])$. We will assume this for the moment.

We write $f \sim \sum_{n=-\infty}^{\infty} a_n e^{inx}$ where $\sim$ denotes (formally) the Fourier series of f, where

$$a_n = \langle f, e_n \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} \mathrm{d}x$$

termed the n^{th} Fourier coefficient .

Note

The Fourier coefficients are orthogonal;

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{inx} e^{-imx} \mathrm{d}x = \frac{1}{2\pi} \frac{1}{i(m-n)} e^{i(m-n)x} \Big|_{-\pi}^{\pi}$$

equal to 1 if $n = m$ and equal to zero otherwise.

Theorem 3.10 ($\mathcal{L}^2$-theory of Fourier series)
Assume $f \in \mathcal{L}^2([-\pi, \pi])$. Then

1. The "classical" Parseval identity

$$\sum_{n=-\infty}^{\infty} |a_n|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 \mathrm{d}x$$

   holds.

2. As $N \to \infty$, we have

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |S_N f(x) - f(x)|^2 \mathrm{d}x \to 0 \tag{3.32}$$

   where $S_N f(x) = \sum_{|n| \le N} a_n e^{inx}$, $a_n = \langle f, e_n \rangle$.

where (3.32) is $\mathcal{L}^2$-convergence of **Fourier series**. (3.32) is a direct consequence of our general reslut modulo showing that $\{e^{inx}\}_{n\in\mathbb{Z}}$ is a Hilbert basis.[8]

First, we note that $\mathcal{L}^2([-\pi, \pi]) \subset \mathcal{L}^1([-\pi, \pi])$ by Cauchy-Schwarz since $m([-\pi, \pi]) = 1 < \infty$. Indeed,

$$\int_{-\pi}^{\pi} |f| \mathrm{d}x \le \left(\int_{-\pi}^{\pi} |f|^2 \mathrm{d}x \right)^{\frac{1}{2}} (2\pi)^{\frac{1}{2}}.$$

We start with the following.

Theorem 3.11
Suppose $f \in \mathcal{L}^1([-\pi, \pi])$. Then

1. If $a_n = 0$ for all $n \Rightarrow f(x) = 0$ almost everywhere.
2. $\sum_{n=-\infty}^{\infty} a_n r^{|n|} e^{inx} \to f(x)$ for a.e. x as $r \to 1^-$, termed a.e. Abel summability.

Proof. Suppose we know (2), then (1) is an immediate consequence.[9] To prove (2), we write

$$\lim_{r\to 1^-} \lim_{N\to\infty} \sum_{|n|\le N} a_n r^{|n|} e^{inx} = f(x)$$

[8] We have a unitary equivalence between ℓ^2 and $\mathcal{L}^2([-\pi, \pi])$

[9] Pointwise convergence almost everywhere is a Fields medal result, which we won't deal with in this course

for a.e.$x \in [-\pi, \pi]$.

The first step is to understand

$$\lim_{N\to\infty} \sum_{|n|\le N} a_n r^{|n|} e^{inx} = \sum_{n=-\infty}^{\infty} r^{|n|} e^{inx}$$
$$= 1 + \sum_{n=1}^{\infty} \left(re^{ix}\right)^n + \sum_{n=1}^{\infty} \left(re^{-ix}\right)^n$$

The key point here is that $|re^{ix}| = r < 1$. So $\sum_{n=1}^{\infty} \left(re^{ix}\right)^n$ is a convergent geometric series. Similar argument for the sum $\sum_{n=1}^{\infty} \left(r^{-1}e^{-ix}\right)^n$. An easy calculation using geometric series (exercise) gives the following explicit formula.

$$\sum_{n=-\infty}^{\infty} r^{|n|} e^{inx} = \frac{1-r^2}{1-2r\cos(x)+r^2}$$

where

$$\mathrm{P}_r(x) = \frac{1-r^2}{1-2r\cos(x)+r^2}$$

is the **Poisson kernel** function disk.

Hint

Take $\sum_{n=0}^{\infty} z^n = (1-z)^{-1}$ for $|z| < 1$ and take complex conjugate.

Note

For $0 < r < 1$, $\mathrm{P}_r(x) \in \mathcal{C}^\infty([-\pi, \pi])$ with $\mathrm{P}_r(-\pi) = \mathrm{P}_r(\pi)$.

For the second step, write $\sum_{n=-\infty}^{\infty} a_n r^{|n|} e^{inx}$ in terms of the Poisson kernel. We argue 'formally' for the moment:

$$f(x) \sim \sum_{n=-\infty}^{\infty} a_n e^{inx} f(x-y) \sim \sum_{n=-\infty}^{\infty} a_n e^{in(x-y)}$$

Integrating $f(x-y)$ against the Poisson kernel, we get

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-y)\mathrm{P}_r(y)\mathrm{d}y &\sim \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\sum_{n=-\infty}^{\infty} a_n e^{in(x-y)}\right)\left(\sum_{m=-\infty}^{\infty} r^{|m|}e^{imy}\right)\\ &\sim \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} a_n \frac{r^{|m|}}{2\pi}\int_{-\pi}^{\pi} e^{in(x-y)}e^{imy}\mathrm{d}y\\ &\sim \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} a_n r^{|m|}e^{inx}\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(m-n)y}\mathrm{d}y\\ &\sim \sum_{n=-\infty}^{\infty} a_n r^{|n|}e^{inx}\end{aligned}$$

since if $m \neq n$, this is zero, so we pick up a single sum. Taking the Poisson kernel with the Fourier expansion induces orthogonality conditions.

Claim

For every $x \in [-\pi, \pi]$ and for $0 < r < 1$ and $f(x+2\pi) = f(x)$

$$\begin{aligned}\sum_{n=-\infty}^{\infty} a_n r^{|n|}e^{inx} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-y)\mathrm{P}_r(y)\mathrm{d}y \\ &= (f * \mathrm{P}_r)(x)\end{aligned} \tag{3.33}$$

where $*$ in $(f * \mathrm{P}_r)$ denotes convolution.

Proof. By dominated convergence,

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-y)\mathrm{P}_r(y)\mathrm{d}y &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-y)\left(\sum_{n=-\infty}^{\infty} r^{|n|}e^{iny}\right)\mathrm{d}y\\ &= \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} r^{|n|}\left(\int_{-\pi}^{\pi} f(x-y)e^{iny}\mathrm{d}y\right)\end{aligned}$$

since $\sum_{n=-\infty}^{\infty} r^{|n|}e^{iny}$'s converges absolutely and uniformly for $y \in [-\pi, \pi]$. Using translation invariance of the Lebesgue measure, and taking $\mathrm{d}y = \mathrm{d}(y-x)$, we obtain

$$\begin{aligned}\int_{-\pi}^{\pi} f(x-y)e^{iny}\mathrm{d}y &= \left(\int_{-\pi}^{\pi} f(x-y)e^{in(y-x)}\mathrm{d}y\right)e^{inx}\\ &= \left(\int_{-\pi}^{\pi} f(y)e^{-iny}\mathrm{d}y\right)e^{inx}\\ &= a_n e^{inx}\end{aligned}$$

■

The last part of the theorem amounts to proving $\lim_{r\to 1^-}(f * \mathrm{P}_r)(x) = f(x)$ for a.e. $x \in [-\pi,\pi]$. We have to show that P_r for $0 < r < 1$ is an approximation to the identity. This will prove $(e^{inx})_{n\in\mathbb{Z}}$ is a Hilbert basis. ■

Approximations of the Identity -good kernels

Claim

$(e^{in\theta})_{n\in\mathbb{Z}}$ is a Hilbert basis for $\mathcal{L}^2([-\pi,\pi])$. To do this, we need to show that

$$(\mathrm{P}_r * f)(x) \xrightarrow{r\to 1^-} f(x)$$

for a.e. $x \in ([-\pi,\pi])$. Here

$$\mathrm{P}_r(x) = \frac{1-r^2}{1-2r\cos(x)+r^2}\mathbf{1}_{[|x|\le\pi]}.$$

We first define the

Definition 3.12 (Approximations to the identity)

We consider a family of functions $(K_\delta)_{\delta>0}$. In the Poisson case, $\delta = 1-r$. We consider convolutions

$$K_\delta * f(x) = \int_{\mathbb{R}^d} f(x-y)K_\delta(y)\mathrm{d}y \tag{3.34}$$

The idea: $K_\delta \to \delta_0$ as $\delta \to 0^+$. We would like the following basic properties

(1) $\int_{\mathbb{R}^d} K_\delta(x)\mathrm{d}x = 1$ for all $\delta > 0$

(2) $\int_{\mathbb{R}^d} |K_\delta(x)|\mathrm{d}x \le A < \infty$. for all $r > 0$.

(3) For fixed $\eta \ne 0$, $\int_{|x|\ge\eta} |K_\delta(x)|\mathrm{d}x \to 0$ as $\delta \to 0^+$.

In the Poisson case, with $\delta = 1-r$, it is easy to see that (1) − (3) are satisfied. For nice functions (*e.g.* $f \in \mathcal{C}^0(\mathbb{R}^d)\cap\mathcal{L}^\infty(\mathbb{R}^d)$; continuous and bounded by M), one can easily check that $(f * K_\delta)(x) \to f(x)$ as $\delta \to 0$ for all δ.

We have the basic identity:

$$\begin{aligned}(f * K_\delta)(x) - f(x) &= \int_{\mathbb{R}^d} f(x-y)K_\delta(y)\mathrm{d}y - f(x)\\ &\stackrel{!}{=} \int_{\mathbb{R}^d} [f(x-y) - f(x)]K_\delta(y)\mathrm{d}y\end{aligned}$$

since $\int_{\mathbb{R}^d} K_\delta(y)\mathrm{d}y = 1$ for all $\delta > 0$. We break into two case

$$= \int_{|y|\le\eta} [f(x-y) - f(x)]K_\delta(y)\mathrm{d}y + \int_{|y|>\eta} [f(x-y) - f(x)]K_\delta(y)\mathrm{d}y$$

and now the term involving $\int_{|y|>\eta} \cdots$ is

$$\int_{|y|>\eta} |f(x-y) - f(x)||K_\delta(y)|\mathrm{d}y \le 2M \int_{|y|>\eta} |K_\delta(y)|\mathrm{d}y \to 0$$

as $\delta \to 0^+$ by property (3). Here, we have $\sup_{x\in\mathbb{R}^d} |f(x)| \le M < \infty$.

For the first term, we use continuity of f at x. Given any ε, can find η such that

$$|f(x-y) - f(x)| < \varepsilon \quad \text{ if } |y| < \eta$$

Thus, the integral part $\int_{|y|\le\eta} \cdots$ can be bounded by

$$\begin{aligned} &\le \int_{\mathbb{R}^d} |f(x-y) - f(x)||K_\delta(y)|\mathrm{d}y \\ &< \varepsilon \int_{\mathbb{R}^d} |K_\delta(y)|\mathrm{d}y \\ &\le A\varepsilon \end{aligned}$$

for all $\delta > 0$ by proposition (2).

We need to deal with functions $f \in \mathcal{L}^1(\mathbb{R}^d)$. To do this, we have to strengthen properties (1)-(3) for (K_δ) slightly, but sot that we still satisfied for all approximations of the identity of interest (*e.g.* Poisson, Heat, Fejer, etc.)

We reformulate the properties in a slightly different manner

- (1′) $\int_{\mathbb{R}^d} K_\delta(x)\mathrm{d}x = 1$ for all $\delta > 0$
- (2′) $|K_\delta(x)| \le A\delta^{-d}$ for all $\delta > 0$ ($|x|$ close to zero).
- (3′) $|K_\delta(x)| \le \frac{A\delta}{|x|^{d+1}}$ for all $\delta > 0$ (for $|x|$ far from zero).

First, we claim that $(2') - (3') \Rightarrow (2) - (3)$. Now

$$\begin{aligned}\int_{\mathbb{R}^d} |K_\delta(x)|\mathrm{d}x &= \int_{|x|\leq\delta} |K_\delta(x)|\mathrm{d}x + \int_{|x|>\delta} |K_\delta(x)|\mathrm{d}x \\ &\leq A\delta^{-d} \int_{|x|\leq\delta} \mathrm{d}x + \int_{|x|\ \delta} \frac{A\delta}{|x|^{d+1}}\mathrm{d}x \\ &\leq A + A'\delta \int_{r>\delta} \frac{r^{d-1}}{r^{d+1}}\mathrm{d}r \\ &= A + A\delta \int_r^\infty \frac{\mathrm{d}r}{r^2} \\ &= A + A''\frac{\delta}{\delta} < \infty\end{aligned}$$

using polar variables (one could also use a dyadic decomposition). To see that property (3) holds,

$$\begin{aligned}\int_{|x|>\eta} |K_\delta(x)|\mathrm{d}x &\leq A\delta \left(\int_{|x|>\eta} \frac{\mathrm{d}x}{|x|^{d+1}}\right) \\ &\leq \frac{A'r}{\eta} \to 0\end{aligned}$$

as $\delta \to 0^+$.

Poisson kernels

(1) is easy to check ($\delta = 1 - r$), for $0 < r < 1$, $\frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{P}_r(x)\mathrm{d}x = 1$. Writing explicitly the Poisson kernel,

$$\mathrm{P}_r(x) = \sum_{n\geq\mathbb{Z}} r^{|n|}e^{inx}$$

implies that

$$\begin{aligned}&\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{P}_r(x)\mathrm{d}x \\ &= \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} r^{|n|}\int_{-\pi}^{\pi} e^{inx}\mathrm{d}x = 1\end{aligned}$$

since $\int_{-\pi}^{\pi} e^{inx}\mathrm{d}x = 2\pi\mathbf{1}_{[\eta=0]}$. Now, for $(2')$, if $|x| \leq \pi$

$$\begin{aligned}\mathrm{P}_r(x) &= \frac{1-r^2}{1-2r\cos(x)+r^2} \\ &= \frac{(1-r)(1+r)}{1-2r\cos(x)+r^2}\end{aligned}$$

and for $x = 0$, the best case offender, we get

$$\frac{(1-r)(1+r)}{(1-r)^2} = \frac{1+r}{1-r} = \frac{2-\delta}{\delta}.$$

For (3′), $|x| > \eta$ for $\eta \neq 0$ and $\eta \in [-\pi, \pi]$, one can check (exercise) that

$$|\mathrm{P}_r(x)| \leq C_0\delta \quad \text{for} |x| > \eta$$

since the quadratic polynomial in the denominator will be uniformly bounded away from zero.

We get an approximation of $\mathcal{L}^p$ function in terms of smooth functions.

To show $f * K_\delta \to f$ a.e. for $f \in \mathcal{L}^1(\mathbb{R}^d)$ and $(K_\delta)_{\delta>0}$ – an approximation of the identity, we will need the following generalization of the Fundamental Theorem of Calculus.

Theorem 3.13 (Lebesgue Differentiation Theorem)

Given $f \in \mathcal{L}^1(\mathbb{R}^d)$ (or more generally $\mathcal{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$). Then, for a.e. $x \in \mathbb{R}^d$,

$$\lim_{\substack{m(B)\to 0 \\ x\in B}} \frac{1}{m(B)} \int_B f(y)dy = f(x).$$

where B is the Euclidian ball.

We will not prove this at the moment.

Definition 3.14 (Lebesgue set)

The Lebesgue set of f ($\mathrm{Leb}(f)$) is the set of $x \in \mathbb{R}^d$ for which the Lebesgue differentiation theorem holds, that is

$$\lim_{\substack{m(B)\to 0 \\ x\in B}} \frac{1}{m(B)} \int_B |f(y) - f(x)|dy = 0$$

$\mathrm{Leb}(f)$ is of full measure, *i.e.* $m((\mathrm{Leb}(f))^{\complement}) = 0$.

There is the following

Corollary 3.15

Assume $f \in \mathcal{L}^1$ (again, $f \in \mathcal{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$ will suffice). Then, a.e. $x \in \mathbb{R}^d$ is in $\mathrm{Leb}(f)$, *i.e.* $m((\mathrm{Leb}(f))^{\complement}) = 0$.

Proof. Let $r \in \mathbb{Q}$ and apply Lebesgue differentiation theorem to the function $g(y) =$

$|f(y)-r|$. Then, by Lebesgue, $\exists E_r \in \mathcal{M}$ with $m(E_r)=0$ such that for $x \notin E_r$

$$\lim_{\substack{m(B)\to 0\\ x\notin E_r}} \int_B |f(y)-r|dy = |f(x)-r|.$$

Now, let $E=\bigcup_{r\in\mathbb{Q}} E_r$. Clearly, $m(E)=0$. Suppose $\bar{x}\notin E$ and that $f(\bar{x})<\infty$. Then, for any $\varepsilon>0$, there exists $r\in\mathbb{Q}$ such that $|f(\bar{x})-r|<\varepsilon$. This implies that

$$\begin{aligned}\frac{1}{m(B)}\int_B |f(y)-f(\bar{x})|dy \\ &\leq \frac{1}{m(B)}\int_B |f(y)-r|dy+|f(\bar{x})-r| \\ &= 2|f(\bar{x})-r|<2\varepsilon\end{aligned}$$

by Lebesgue. ■

We will also use the following **absolute continuity** result for Lebesgue integral.

Lemma 3.16

Assume $f\in\mathcal{L}^1(\mathbb{R}^d)$. Then for any $\varepsilon>0$, there exists $\delta=\delta(\varepsilon)>0$ such that

$$\int_E |f|<\varepsilon$$

provided $m(E)<\delta$.

Proof. Left as an exercise. ■

Theorem 3.17

Given $f\in\mathcal{L}^1(\mathbb{R}^d)$ and $(K_\delta)_{\delta>0}$ as above. Then, $f*K_\delta(x)\xrightarrow{\delta\to 0^+} f(x)$ for all $x\in\text{Leb}(f)$ (in particular, for a.e. x).

Proof. Proceeds as before: we use that $\int_{\mathbb{R}^d} K_\delta(x)dx=1$ for all $\delta>0$ to get that

$$\begin{aligned}|(f*K_\delta)(x)-f(x)| &= \left|\int_{\mathbb{R}^d}(f(x-y)-f(x))K_\delta(y)dy\right| \\ &\leq \int_{\mathbb{R}^d}|f(x-y)-f(x)||K_\delta(y)|dy \\ &\leq \int_{|y|<\delta}|f(x-y)-f(x)|||K_\delta(y)|dy \qquad (3.35a)\\ &\quad+\int_{|y|\geq\delta}|f(x-y)-f(x)|||K_\delta(y)|dy \qquad (3.35b)\end{aligned}$$

To estimate (3.35a), we recall that $|K_\delta(y)| \leq \delta\delta^{-d} \leq c/m(B(\delta))$ and thus we have (3.35a)

$$\leq \frac{c}{\delta^d}\int_{|y|<\delta}|f(x-y)-f(x)|dy.$$

Now, we need the following.

Claim

Given $f \in \mathcal{L}^1(\mathbb{R}^d)$ and $x \in \text{Leb}(f)$, then consider

$$A(\delta) = \frac{1}{\delta^d}\int_{|y|\leq\delta}|f(x-y)-f(x)|dy$$

where $A : \mathbb{R}^+ \to \mathbb{R}^+$. The function $A(\delta)$ has the folllowing important properties.

1. $A \in \mathcal{C}^0(\mathbb{R}^+)$
2. $A \in \mathcal{L}^\infty(\mathbb{R}^+)$, that is uniformly bounded: $A(\delta) \leq M$ for all $\delta > 0$.
3. $\lim_{\delta\to 0^+} A(\delta) = 0$.

Proof.

1. $A \in \mathcal{C}^0(\mathbb{R}^+)$ follows from absolute continuity of Lebesgue integral
2. If $0 \leq \delta \leq 1$, since $A(\delta) \in \mathcal{C}^0(\mathbb{R}^+)$ with $\lim_{\delta\to 0^+} m(B) = 0$ implies $A(\delta) \leq M$ for all $\delta \in [0,1]$. When $\delta > 1$, $A(\delta) \leq c\left(\|f\|_{\mathcal{L}^1} + |f(x)|\right) < \infty$ for $x \in \text{Leb}(x)$.
3. This is an application of the corollary (3.15) to Lebesgue differentiation, since $m(B(\delta)) \sim c_d\delta^d$

■

We can now finally finish the proof of the theorem. We split $\int_R^d > |f(x-y)-f(x)||K_\delta(y)|dy$ into

$$\begin{aligned}\int_R^d > |f(x-y)-f(x)||K_\delta(y)|dy &= \int_{|y|\leq\delta}(\cdots) + \sum_{k=0}^{\infty}\int_{2^k\delta\leq|y|\leq 2^{k+1}\delta}\int(\cdots)\\ &= Ac(\delta) + \sum_{k=0}^{\infty}\int_{2^k\delta\leq|y|\leq 2^{k+1}\delta}\int(\cdots)\end{aligned}$$

and by the lemma, $A(\delta) \to 0$ as $\delta \to 0^+$. For the second term, using the decay assumption on K_δ

$$\begin{aligned}
&\leq \frac{c\delta}{(2^k\delta)^{k+1}} \int_{|y|\leq 2^{k+1}\delta} |f(x-y)-f(x)|dy \\
&\leq \frac{c'}{2^k(2^{k+1}\delta)^d} \int_{|y|\leq 2^{k+1}\delta} |f(x-y)-f(x)|dy \\
&\leq c2^{-k}A(2^{k+1}\delta).
\end{aligned}$$

The upshot is that

$$|f * K_\delta(x) - f(x)| \leq cA(\delta) + c' \sum_{k=0}^{\infty} 2^{-k}A(2^{k+1}\delta)$$

for $x \in \mathrm{Leb}(f)$.

We have shown that

$$|f * K_\delta(x) - f(x)| \leq c_1 A(\delta) + c_2 \sum_{k=0}^{\infty} 2^{-k}A(2^{k+1}\delta). \tag{3.36}$$

Given $\varepsilon > 0$, by choosing $N > 0$ large, we can arrange that $\sum_{k\geq N} 2^{-k} < \varepsilon$. Since $A(r) \to 0$ as $r \to 0^+$, by choosing $\delta > 0$ small enough, we can arrange that $A(2^k\delta) < \frac{\varepsilon}{N}$ for $k = 0, 1, \ldots, N-1$. This implies that the RHS of (3.36) is less than or equal to

$$\begin{aligned}
&c_1'\varepsilon + c_2' \sum_{k=0}^{N-1} \frac{\varepsilon}{N} + c_2 \sum_{k=N}^{\infty} 2^{-k}A(2^{k+1}\delta) \\
&\leq c_3'\varepsilon + M\left(\sum_{k=N}^{\infty} 2^{-k}\right) < \varepsilon
\end{aligned}$$

by picking N large at the outset. Here, $M = \|A(\delta)\|_{\mathcal{L}^\infty(\mathbb{R}^+} < \infty$. Therefore, for $x \in \mathrm{Leb}(A)$, we have

$$|f * K_\delta(x) - f(x)| < c_4\varepsilon \Rightarrow (f * K_\delta)(x) \xrightarrow{\delta\to 0^+} f(x)$$

for a.e. $x \in \mathbb{R}^d$ ■

Corollary 3.18

Applying this result to the **Poisson kernel,**

$$\mathrm{P}_r(y) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{iny} \mathbf{1}_{[y \in [-\pi,\pi]]}$$

we get a.e. Abel convergence,

$$\lim_{r \to 1^-} \sum_{n=-\infty}^{\infty} r^n a_n e^{iny} = f(x)$$

a.e. x. As a consequence, we get that $(e^{inx})_{n\in\mathbb{Z}}$ is a Hilbert basis for $\mathcal{L}^2([-\pi,\pi])$.

3.4 Application of approximations to the identity to complex analysis and PDE

Consider $\mathbb{R}/2\pi\mathbb{Z}$ can be identified to the unit circle $S^1 = \{z \in \mathbb{C} | |z| = 1\}$ with the usual parametrization $[-\pi,\pi] \ni \theta \to e^{i\theta}$.

Consider a simplified variant of the Dirichlet problem: solve the following boundary value problem:

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

which is the Laplacian and the boundary value problem

$$\begin{cases} \Delta u = 0 \text{ in } D = \{(x,y) \in \mathbb{R}^2 | x^2 + y^2 < 1\} \\ u\big|_{\partial D} = f \in \mathcal{L}^2 \text{ a.e.} \end{cases} \tag{3.37}$$

We can think of $f \in \mathcal{L}^2([-\pi,\pi])$, $\partial D = S^1 \cong \mathbb{R}/2\pi\mathbb{Z}$. Here, $u|_{\partial D} = \lim_{r\to 1^-} u(re^{i\theta})$

Remark

The regularized **heat-equation** $(\partial_t - \Delta)u = 0$ and the **wave-equation** $(\partial_t^2 - \Delta)u = 0$. If u is stationary (independent of time), in both cases, $\Delta u = 0$. The questions that could be asked is as to whether there exists a solution and whether it is unique. The answer to both questions is yes, and (3.37) has a unique solution that can be written explicitly in terms of $\mathrm{P}_r(y)$.

The motivation here is the following: recall $\mathrm{P}_r(y) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{iny}$ for $0 \le r \le 1$. Consider the convolution $(f * K_\delta)(x)$ where f is 2π periodic, that is $f(y + 2\pi k) = f(y)$ for all $k \in \mathbb{Z}$.

Consider

$$(f * \mathrm{P}_r)(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-y)\mathrm{P}_r(y)\mathrm{d}y$$
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{P}_r(x-y)f(y)\mathrm{d}y$$

by invariance of $\mathrm{d}y$ under translation. Write $x = \theta \in [-\pi, \pi]$. Let

$$u(re^{i\theta}) = (f * \mathrm{P}_r)(\theta)$$
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{P}_r(\theta - y)f(y)\mathrm{d}y \tag{3.38}$$

where y is the incoming variable and $(r, \theta) \in D$ is the outgoing variable. The kernel is a function of both the incoming and the outgoing variable. We write $\mathrm{P}_r(re^{i\theta}, y) = \mathrm{P}_r(\theta - y)$, so (3.38) becomes

$$u(re^{i\theta}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{P}_r(re^{i\theta}, y)f(y)\mathrm{d}y$$

By the previous argument,

$$\lim_{r\to 1^-} u(re^{i\theta}) = f(\theta)$$

for a.e. $\theta \in [-\pi, \pi]$. We thus need to consider $u(re^{i\theta})$. Write $z = re^{i\theta}$ for $0 \leq r < 1$. Then

$$\mathrm{P}(z, y) = \mathrm{P}_r(\theta - y)$$
$$= \sum_{n=-\infty}^{\infty} r^{|n|}e^{in(\theta-y)}$$
$$= \sum_{n=0}^{\infty} r^n e^{in(\theta-y)} + \sum_{n=1}^{\infty} r^n e^{-in(\theta-y)}$$
$$= \sum_{n=0}^{\infty} z^n e^{iny} + \sum_{n=1}^{\infty} \overline{z}^n e^{iny}$$

for $|z| < 1$ as we have an absolutely uniformly convergent series. We have

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = 4\frac{\partial}{\partial z}\frac{\partial}{\partial \overline{z}}$$

as

$$\frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - \frac{\partial}{\partial y}\right)$$
$$\frac{\partial}{\partial \overline{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right)$$

and

$$\frac{1}{4}\Delta u = \frac{\partial}{\partial z}\frac{\partial}{\partial \overline{z}}\left(\sum_{n=0}^{\infty} z^n e^{iny} + \sum_{n=1}^{\infty} \overline{z}^n e^{iny}\right) = 0$$

Then, by Dominated convergence theorem (exercise), we can differentiate under the integral sign to get

$$\frac{1}{4}\Delta u(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{\partial}{\partial z}\frac{\partial}{\partial \overline{z}}\mathrm{P}_(z,y)\right) f(y)dy = 0$$

since its is dominated by an $\mathcal{L}^1$ function, in this case zero. This can be viewed by rewritting the derivative in terms of limits.

Holomorphic $\overline{\partial}$ problem

Consider a disk, with $u|_{d\Omega} = f$. The $\overline{\partial}$ problem goes as follow:

$$\begin{cases} \overline{\partial} u = 0 \\ u|_{\partial D} = f \in \mathcal{L}^2, \end{cases} \tag{3.39}$$

where $u|_{\partial D} = \lim_{r \to 1^-} u(re^{i\theta})$ and where $\overline{\partial} = \frac{1}{2}(\partial_x + i\partial_y)$ *i.e.* find a holomorphic function $u(z)$ in D with prescribed boundary values.

(3.39) cannot be solved for arbitrary $f \in \mathcal{L}^2(\delta^{-1})$, unlike Dirichlet. For example, $f(\theta) = e^{-i\theta}$, with $z = re^{i\theta}$ and $u(re^{i\theta} = r^{-1}e^{-i\theta}$. Then

$$f(\theta) \sim \sum_{n=0}^{\infty} a_n e^{in\theta} + \sum_{n=-1}^{-\infty} a_n e^{in\theta}$$

If we have Hardy functions, we can solve the problem. To solve this problem, we have already seen that Fourier coefficients of $f, (a_n)_{n<0}$ are problematic.

Definition 3.19
We say that $f \in \mathcal{L}^2([-\pi, \pi])$ is in the $\mathcal{L}^2$-Hardy space, $\mathcal{H}^2(D)$, provided that $f \sim \sum_{n=0}^{\infty} a_n e^{in\theta}$, namely all negative Fourier coefficients are zero.

Since harmonic functions are just real parts of holomorphic functions on D, we are motivated by the Dirichlet problem. The Poisson kernel is given by for $0 \leq r < 1$

$$\begin{aligned}
\mathrm{P}_r(\theta) &= \sum_{n=0}^{\infty}(re^{i\theta})^n + \sum_{n=1}^{\infty}(re^{-i\theta})^n \\
&= \sum_{n=0}^{\infty} z^n + \sum_{i=1}^{n} \overline{z}^n \\
&= \sum_{n=0}^{\infty}(r^n e^{in\theta}) + \sum_{n=1}^{\infty}(r^n e^{-in\theta})\mathrm{d}y
\end{aligned}$$

for $|z| < 1$. Now, let $f \in \mathcal{H}^2(D)$ and consider

$$\begin{aligned}
&\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{P}_r(\theta - y)f(y)\mathrm{d}y \\
&\qquad = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{n=0}^{\infty} ir^n e^{in(\theta-y)}f(y)\mathrm{d}y + \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{n=1}^{\infty} ir^n e^{-in(\theta-y)}f(y)\mathrm{d}y
\end{aligned}$$

and by DCT,

$$= \sum_{n=0}^{\infty}\frac{1}{2\pi}(re^{i\theta})^n\int_{-\pi}^{\pi} e^{-iny}f(y)\mathrm{d}y + \sum_{n=1}^{\infty}\frac{1}{2\pi}(re^{i\theta})^n\int_{-\pi}^{\pi} e^{iny}f(y)\mathrm{d}y$$

and the righmost term is 0 since $f \in \mathcal{H}^2(D)$. Therefore, the convolution is just

$$\begin{aligned}
&\frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{n=0}^{\infty}(re^{i\theta})^n e^{iny}f(y)\mathrm{d}y \\
&\qquad = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{n=0}^{\infty} z^n e^{iny}f(y)\mathrm{d}y
\end{aligned}$$

for $z = re^{i\theta}$. Here, $C(z, y) \coloneqq \sum_{n=0}^{\infty} z^n e^{iny}$ for $(z, y) \in D \times \partial D$ is called the **Cauchy kernel.**

$$\sum_{n=0}^{\infty}\left(\frac{z}{e^{iy}}\right)^n = \frac{1}{1 - \frac{z}{e^{iy}}} = \frac{e^{iy}}{e^{iy} - z}$$

where $|z| < 1, e^{iy} \leftarrow \partial D$.

Theorem 3.20

Given $f \in \mathcal{H}^2(D)$, there is a unique solution $\varphi(z)$ to $\overline{\partial}$ problem in D given by

$$u(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi} C(z, y)f(y)\mathrm{d}y.$$

Write $f(y) = F(e^{iy})$.

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} C(z,y)F(e^{iy})\mathrm{d}y = \frac{1}{2\pi}\int_{-\pi}^{p} i\frac{e^{iy}}{e^{iy}-z}F(e^{iy})\mathrm{d}y$$

Let $w = e^{iy}, dw = ie^{iy}\mathrm{d}y$, we want an i factor to do the contour integral to use the Cauchy integral formula. We have the complex contour

$$\frac{1}{2\pi i}\oint_{S^1}\frac{F(w)}{w-z}\mathrm{d}w$$

3.5 Closed subspaces of Hilbert spaces

The main point about a Hilbert space is **completeness**. Given $\{f_n\}_{n=1}^{\infty} \in H$ Cauchy, $f_n \to f$ as $n \to \infty$ with $f \in H$. As we know, this completeness is in general false. Let RI([0,1]) denote the space of Riemann integrable functions on [0,1]. Given $f, g \in \mathsf{RI}$([0,1]) $\subseteq \mathcal{L}^2([0,1])$ and $\alpha f + \beta g \in \mathsf{RI}$([0,1]). So RI([0,1]) is a linear subspace. However, given $(f_n)_{n=1}^{\infty} \in \mathsf{RI}$([0,1]) with $f_n \xrightarrow{\mathcal{L}^2} f$ as $n \to \infty$. It is not true that $f \in$ RI in general. This motivates the following

Definition 3.21

A linear subspace $S \subseteq H$ is closed provided H is complete, *i.e.* given $(f_n)_{n=1}^{\infty} \subseteq S$ with $f_n \to f$ as $n \to \infty$ in H implies $f \in S$.

The following is immediate

Proposition 3.22

Any closed subspace $S \subseteq H$ is itself a Hilbert with the induced inner product inherited from H.

A trivial example is the following

Example 3.3

If $\dim(H) < \infty$, then all subspaces $S \subseteq H$ are closed.

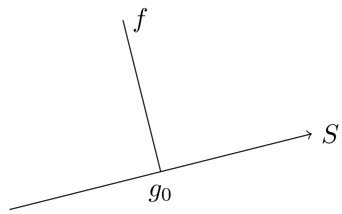
For orthogonal projections, closed subspaces $S \subseteq \mathcal{H}$ mimick finite dimensions. In particular, one has the notion of an orthogonal projection.

Lemma 3.23

Given $S \subseteq \mathcal{H}$ closed subspace of a Hilbert space $\mathcal{H}$ and $f \in \mathcal{H}$. Then

1. There exists a unique $g_0 \in S$ closest to $f \in \mathcal{H}$ in the sense that

$$\|f - g_0\| = \inf_{g \in S}\|f - g\|$$



2. $(f - g_0) \perp S$, namely $\langle f - g_0, g\rangle = 0$ for all $g \in S$.

Remark

The main consequence of this lemma is the existence of an orthogonal decomposition $\mathcal{H} = S \oplus S^{\perp}$, where $S, S^{\perp}$ are both closed.

Proof. If $f \in S$, we are done. Suppose $\notin S$, then $d = \inf_{g \in S} \|f - g\| > 0$ since S is closed. Let $\{g_n\}_{n=1}^{\infty}$ be a sequence in S with

$$\lim_{n \to \infty} \|f - g_n\| = d > 0$$

Claim

$\{g_n\}_{n=1}^{\infty} \subset S$ is Cauchy.

Proof. We have to write $\|g_n - g_m\|$ in terms of $\|f - g_n\|$ and $\|f - g_m\|$; we use parallelogram law

$$\|A + B\|^2 + \|A - B\|^2 = 2\left(\|A\|^2 + \|B\|^2\right)$$

for $A, B \in \mathcal{H}$. Apply the parallelogram law with $A = f - g_n, B = f - g_m$. We get

$$\|2f - (g_n + g_m)\|^2 + \|g_n - g_m\|^2 = 2\left(\|f - g_n\|^2 + \|f - g_m\|^2\right) \tag{3.40}$$

Use the fact that

$$\|2f - (g_n + g_m)\|^2 = 4\left\|f - \frac{g_n + g_m}{2}\right\|^2 \geq 4d^2$$

since $(g_n + g_m)/2 \in S$. This implies that

$$\|g_n - g_m\|^2 = 2\left(\|f - g_n\|^2 + \|f - g_m\|^2\right) - \|2f - (g_n + g_m)\|^2 - 4d^2$$

We know that $\|f - g_n\| \searrow d$ as $n \to \infty$ and $\|f - g_m\| \searrow d$ as $m \to \infty$ by assumption. Therefore $\{g_n\}_{n=1}^{\infty}$ is Cauchy. ■

Since $S \subseteq \mathcal{H}$ is closed, $\lim_{n\to\infty} g_n$ exists and we call it $g_0 \in S$. Then, $\lim_{n\to\infty} \|f - g_n\| = \|f - g_0\| = d$. We will prove uniqueness at the end.

Now, for the orthogonality, let $g \in S$. We want to show that $\langle f - g_0, g\rangle = 0$. Consider the perturbation $g_0 \mapsto g_0 - \varepsilon g \in S$ for $|\varepsilon| > 0$ small. Since $g_0 \in S$ is a minimizer,

$$\|f - (g_0 - \varepsilon g)\|^2 \geq \|f - g_0\|^2 \tag{3.41}$$

We expand the LHS in (3.41)

$$2\varepsilon \Re\langle f - g_0, g\rangle + \varepsilon^2 \|g\|^2 \geq 0$$

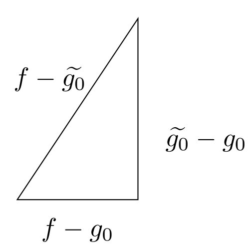
If $\langle f - g_0, g\rangle \lessgtr 0$. Then, taking $\varepsilon \gtrless 0$ sufficiently small gives a contradiction. So the only possibility is $\Re\langle f - g_0, g\rangle = 0$. To deal with $\Im\langle f - g_0, g\rangle$ we make the pertrbabtion $g_0 \mapsto g_0 - i\varepsilon g$ and get that $\Im\langle f - g_0, g\rangle = -$ for all $g \in S$. Finally, to prove uniqueness, we assume that $\widetilde{g_0}$ is another minimizer. Let $g = g_0 - \widetilde{g_0} \in S$. We know

$$\langle f - g_0, g_0 - \widetilde{g_0}\rangle = 0$$

by (2). By Pythagoras,

$$\|f - \widetilde{g_0}\|^2 = \|f - g_0\|^2 + \|\widetilde{g_0} - g_0\|^2$$

■



Proposition 3.24

Let $S \subset H$ be closed and

$$S^\perp = \{f \in \mathcal{H} \mod \langle f, g\rangle = 0 \text{ for all } g \in S\}$$

Then $S^\perp \subset \mathcal{H}$ is a closed subspace and

$$\mathcal{H} = S \oplus S^\perp$$

and $S \cap S^\perp = \{0\}$.

Note

(3.24) means that any $f \in \mathcal{H}$ can be written uniquely in the form $f = g + h$ with $g \in S$ and $h \in S^\perp$.

Remark

Since S and $S^\perp$ are themselves Hilbert spaces, we can iterate this procedure to refine this decomposition.

Proof. The fact that $S^\perp \subset \mathcal{H}$ is linear is clear. To see that it is closed, we use Cauchy-Schwarz: let $\{f_n\}_{n=1}^\infty \in S^\perp$ with $\|f_n - f\| \to 0$. By assumption, $\langle f_n, g\rangle = 0$ for all $g \in S$. We want to show that the following go to zero:

$$\begin{aligned} |\langle f, g\rangle - \langle f_n, g\rangle| &= |\langle g, f - f_n\rangle| \\ &\leq \|g\| \|f - f_n\| \to 0 \end{aligned}$$

as $n \to \infty$, therefore $f \in S^\perp$.

Suppose we have 2 decompositions, $f = g + h, f = \tilde{g} + \tilde{h}$ where $g, \tilde{g} \in S, f$ and $\tilde{f} \in S^\perp$. Writing

$$S \ni g - \tilde{g} = \tilde{h} - h \in S^\perp$$

and since $S \cap S^\perp = \{0\}$ we conclude $g = \tilde{g}$ and $h = \tilde{h}$. ■

3.6 Linear transformations

Definition 3.25

1. Given Hilbert spaces, $\mathcal{H}_1, \mathcal{H}_2$, a linear transformation $T : \mathcal{H}_1 \to \mathcal{H}_2$ is a map with

$$T(\alpha f + \beta g) = \alpha T f + \beta T g$$

for all $\alpha, \beta \in \mathbb{C}$ and $f, g \in \mathcal{H}_1$

2. We say that $T : \mathcal{H}_1 \to \mathcal{H}_2$ is **bounded** if $\exists M < \infty$ such that

$$\|Tf\|_{\mathcal{H}_2} \leq M\|f\|_{\mathcal{H}_1} \tag{3.42}$$

   for all $f \in \mathcal{H}_1$.

3. If $T : \mathcal{H}_1 \to \mathcal{H}_2$ is bounded, its norm

$$\|T\| := \inf M$$

   in (3.42). One can compute $\|T\|$ in several ways

Lemma 3.26
Assume $T : \mathcal{H}_1 \to \mathcal{H}_2$ is bounded. Then

$$\|T\| = \sup\{|\langle Tf, g\rangle_{\mathcal{H}^2}| : \|f\|_{\mathcal{H}_1} \leq 1 \text{ and } \|g\|_{\mathcal{H}_2} \leq 1\} \tag{3.43}$$

Proof.

($\Rightarrow$) Assume $\|T\| \leq M$, then

$$\begin{aligned} |\langle Tf, g\rangle_{\mathcal{H}_2}| &\leq \|Tf\|_{\mathcal{H}_2}\|g\|_{\mathcal{H}_2} \\ &\leq M\|f\|_{\mathcal{H}_1}\|g\|_{\mathcal{H}_2} \end{aligned}$$

by Cauchy-Schwarz. If $\|f\| \leq 1, \|g\| \leq 1$, this implies

$$|\langle Tf, g\rangle| \leq M$$

implies the RHS of (3.43) is less than or equal to M.

($\Leftarrow$) Conversely, assume that

$$\sup\{|\langle Tf, g\rangle_{\mathcal{H}^2} \mid |\|f\|_{\mathcal{H}_1} \leq 1 \text{ and } \|g\|_{\mathcal{H}_2} \leq 1\} \leq M$$

It suffices to assume that $f, g \neq 0$ (why?)

Consider $f' = \frac{f}{\|f\|}$ and $g' = \frac{Tf}{\|Tf\|}$. Then, by assumption,

$$|\langle Tf', g'\rangle| \leq M \Leftrightarrow \left\langle \frac{Tf}{\|f\|}, \frac{Tf}{\|f\|} \right\rangle = \frac{\|Tf\|}{\|f\|}$$

and therefore $\|Tf\| \leq M\|f\|$. ■

Definition 3.27

We say that $f : \mathcal{H}_1 \to \mathcal{H}_2$ is continuous provided $\|Tf - Tf_n\| \to 0$ when $\|f - f_n\| \to 0$.

Proposition 3.28

$T : \mathcal{H}_1 \to \mathcal{H}_2$ is continuous if and only if it is bounded

Proof.

($\Rightarrow$) Assume $T : \mathcal{H}_1 \to \mathcal{H}_2$ is bounded. Then

$$\begin{aligned} \|Tf - Tf_n\| &= \|T(f - f_n)\| \\ &\leq \|T\| \cdot \|f - f_n\| \end{aligned}$$

where $\|T\| < \infty$, which implies if $\|f - f_n\| \to 0$, then $\|Tf - Tf_n\| \to 0$

($\Leftarrow$) Assume that $T : \mathcal{H}_1 \to \mathcal{H}_2$ is continuous and suppose not. Then, $\forall\, n > 0$, there exists $f_n \in \mathcal{H}, f_n \neq 0$ with $\|Tf_n\| \geq n\|f_n\|$. Consider the vector

$$g_n = \frac{f_n}{n\|f_n\|} \in \mathcal{H}_1$$

Clearly, $\|g_n\| = n^{-1} \to 0$ as $n \to \infty$. By assumption, $\|Tg_n\| \to 0$ as $n \to \infty$, thus

$$\frac{\|Tf_n\|}{n\|f_n\|} \geq 1$$

but should tend to zero. Contradiction. ∎

3.7 Riesz representation

Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces. We have been discussing bounded linear functionals $T : \mathcal{H}_1 \to \mathcal{H}_2$, meaning that $\|Tf\|_{\mathcal{H}_2} \leq M\|f\|_{\mathcal{H}_1}$ where $M < \infty$. Recall that $\|T\| := \inf M$. The Riesz representation characterizes bounded linear transforms $l : \mathcal{H} \to \mathbb{C}(\mathbb{R})$ where $\mathcal{H}$ is an arbitrary Hilbert space and $\mathbb{C}$ is the simple Hilbert space with norm given by $|\cdot|$. Such linear transformations are called **linear functionals**.

Example 3.4

Fix any vector $g \in \mathcal{H}$ and consider $l : \mathcal{H} \to \mathbb{C}$ given by $l(f) = \langle f, g\rangle$. Clearly, $l : \mathcal{H} \to \mathbb{C}$ is linear and boundedness follows from Cauchy-Schwarz:

$$|l(f)||\langle f, g\rangle| \leq \|g\| \cdot \|f\| < \infty$$

so $l : \mathcal{H} \to \mathbb{C}$ is a **bounded** linear functional

The Riesz theorem says that these are the only bounded linear functionals.

Theorem 3.29 (Riesz representation)

Assume that $l : \mathcal{H} \to \mathbb{C}$ is a continuous linear functional. Then, there exist a unique $g \in \mathcal{H}$ such that

$$l(f) = \langle f, g \rangle$$

and $\|l\| = \|g\|$.

Proof. We consider a particularly useful orthogonal decomposition of $\mathcal{H}$. Consider the null space

$$S = \{f \in \mathcal{H} \mid l(f) = 0\}$$

Claim

1. S is linear since $l(\alpha f_1 + \beta f_2) = \alpha l(f_1) + \beta l(f_2)$.
2. $S \subset \mathcal{H}$ is closed since $l : H \to \mathbb{C}$ is continuous.

Proof. Consider a Cauchy sequence $\{l(f_n)\}_{n=1}^{\infty} \in S, f_n \in \mathcal{H}$ with $|l(f_n) - g| \to 0$ as $n \to \infty$. Since $l : \mathcal{H} \to \mathbb{C}$ is continuous, for the Cauchy limit of f_n, we have $|l(f_n) - l(f)| \to 0$ as $n \to \infty$ is immediate ■

Either $S = \mathcal{H}$ (in this case, we are done and choose $g = 0$) **or** $S \neq \mathcal{H}$. In the latter case, we have an orthogonal decomposition $\mathcal{H} = S \oplus S^{\perp}$.

We choose $h \in S^{\perp}$ with $\|h\| = 1$. Consider the vector $u \in \mathcal{H}$

$$u = l(f)h - l(h)f$$

where $f \in \mathcal{H}, h \in S^{\perp}, \|h\| = 1$. Clearly, $l(u) = 0$ and so $u \in S$. So $\langle u, h \rangle = 0$ if and only if

$$l(f)\|h\|^2 - l(h)\langle f, h \rangle = 0 \tag{3.44}$$

But $\|h\| = 1$ and $l(f) = \langle f, \overline{l(h)} \cdot h \rangle$ therefore $g = \overline{l(h)}h$. Uniqueness is obvious; as for the norm, note that for any linear functional $l(f) = \langle f, g \rangle$ $\|l\| = \|g\|$ since

$$|l(f)| \leq \|f\| \cdot \|g\|$$

by Cauchy-Schwartz, which implies that $\|l\| \leq \|g\|$. But , when $f = g, l(g) = \|g\|^2 = \|g\| \cdot \|g\|$ and the norm of the transformation equals the norm of g in $\mathcal{H}$. ■

One important application of this is the notion of an adjoint (more next week).

Remark

Suppose $\mathcal{H}_0 \subset \mathcal{H}$ is a pre-Hilbert space in the sense that $\overline{\mathcal{H}_0} = \mathcal{H}$ is a Hilbert space. We call $\mathcal{H}$ the completion of $\mathcal{H}_0$.

Let $l_0 : \mathcal{H}_0 \to \mathbb{C}$ be a bounded linear functional with

$$|l_0(f)| \leq M\|f\|, \qquad \forall\, f \in \mathcal{H}_0$$

Then, $l_0 \in \mathcal{H}_0^*$ (notation for bounded linear functional on $\mathcal{H}_0$) has a unique extension to a linear functional $l \in \mathcal{H}^*$ with $|l(f)| \leq M\|f\|$. To construct this extension, we consider the Cauchy sequence $\{l(f_n)\}$ where $\{f_n\} \in \mathcal{H}_0$ with $\|f_n - f\| \to 0 \in \mathcal{H}$.

We define $l(f) \coloneqq \lim_{n\to\infty} l(f_n)$.

3.8 Adjoints

The Riesz theorem allows us to characterize the adjoint of a bounded linear transformation $T : \mathcal{H} \to \mathcal{H}$

Proposition 3.30 (Adjoint of T)

Let $T : \mathcal{H} \to \mathcal{H}$ be a bounded linear transformation and $\mathcal{H}$ a Hilbert space. Then, there exists a unique linear transformation $T^* : \mathcal{H} \to \mathcal{H}$ (termed **adjoint** of T) with

1. $\langle Tf, g\rangle = \langle f, T^*g\rangle$
2. $\|T\| = \|T^*\|$
3. $(T^*)^* = T$

Proof.

1. (Existence) Fix $g \in H$ and consider the following linear functional $l(f) = \langle Tf, g\rangle$ for $f \in \mathcal{H}$. and $l \in \mathcal{H}^*$. Since

$$\begin{aligned} |l(f)| &\leq |\langle Tf, g\rangle| \\ &\leq \|g\| \cdot \|Tf\| \\ &\leq \|g\| \cdot \|T\| \cdot \|f\| \end{aligned}$$

(!) By Riesz, there exists a unique $h \in \mathcal{H}$ such that $l(f) = \langle f, h\rangle$, *i.e.*

$$\langle Tf, g\rangle = \langle f, h\rangle. \tag{3.45}$$

So (3.45) allows us to define $T^* : \mathcal{H} \to \mathcal{H}$ by $T^*g = h$.

2.

$$\begin{aligned}\|T\| &= \sup\{|\langle Tf, g\rangle|; \|f\| \leq 1, \|g\| \leq 1\} \\ &= \sup\{|\langle f, T^*g\rangle|; \|f\| \leq 1, \|g\| \leq 1\} \\ &= \|T^*\|\end{aligned}$$

Think about real symmetric matrices: their adjoints are the matrices themselves.

3.

$$\langle (T^*)^* f, g\rangle = \langle T^* f, T^* g\rangle \tag{3.46}$$

for all f and g if and only if (3.46) holds for all f and g, as one can see by taking complex conjugates and reversing the roles of f and g.

■

3.9 Compact Operators

Compact operators are bounded linear operators $T : \mathcal{H} \to \mathcal{H}$ that most closely resemble finite-dimensional matrices.

Example 3.5

Consider $\mathcal{H} = \mathcal{L}^2([-\pi, \pi]; \mathrm{d}x)$ where $\mathrm{d}x$ denotes the Lebesgue measure. We have the Hilbert basis $\{e^{inx}\}_{n=-\infty}^{\infty}$. Consider the Laplacian $\Delta = \frac{\mathrm{d}^2}{\mathrm{d}x^2}$ on $\mathcal{C}^\infty([-\pi, \pi])$

$$-\Delta(e^{inx}) = \left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^2 e^{inx} = n^2 e^{inx}$$

The functions $\{e^{inx}\}_{n\in\mathbb{Z}}$ play the role of eigenfunctions (or eigenvectors) of the Laplace operator. The eigenfunctions $\{e^{inx}\}_{n\in\mathbb{Z}}$ and the corresponding eigenvalues $\{n^2\}_{n\in\mathbb{Z}}$.

Example 3.6

Consider the ordinary differential operator $P = -\frac{\mathrm{d}^2}{\mathrm{d}x^2} + q(x)$ where $g \in \mathcal{C}^\infty([-\pi, \pi], \mathbb{R})$, where $q(x + 2\pi) = q(x)$ is periodic. This is the 1d Schroedinger operator, or the **Floquet operator**. It turns out there is a direct analogue of Example 3.5 for P, namely there exists a Hilbert basis $\{\varphi_{\lambda_k}\}_{k=1}^{\infty}$ of P with (real) eigenvalues $\{\lambda_k\}_{k=1}^{\infty}$ with

$$P\varphi_{\lambda_k} = \lambda_k \varphi_{\lambda_k}$$

Note

P is highly unbounded in both Examples 3.5 and 3.6, but one show that there exists an operator

$$K : \mathcal{L}^2([-\pi,\pi]) \Rightarrow \mathcal{L}^2([-\pi,\pi])$$

called **Green's operator** (or **parametric** for the approximate inverse) such that $PK = \mathrm{I_d}$. Here, K is a compact operator (nice spectral theory); P itself has a nice spectral decomposition. In particular, we will see that K is an integral operator:

$$K(f) = -\int_{-\pi}^{\pi} K(x,y)f(y)\mathrm{d}y$$

where $K(x,y)$ is a "nice function".

Example 3.7

Consider $P = -\frac{\mathrm{d}^2}{\mathrm{d}x^2} + 1$; then $\varphi_n(x) = e^{inx}$ and

$$P\varphi_n(x) = (n^2+1)\varphi_n(x); n \in \mathbb{Z}$$

To define $K : \mathcal{L}^2 \to \mathcal{L}^2$, we put

$$K\varphi_n(x) = \frac{1}{n^2+1}\varphi_n(x)$$

for $n \in \mathbb{Z}$ and $P \cdot K = \mathrm{I_d}$. One can easily check that $K(x,y) = \sum_{n\in\mathbb{Z}} \frac{e^{in(x-y)}}{n^2+1}$ for $(x,y) \in [-\pi,\pi]\times[-\pi,\pi]$; the eigenvalues of K are

$$\left\{1, \frac{1}{2}, \frac{1}{5}, \ldots\right\}; \quad \frac{1}{n^2+1} \to 0 \text{ as } n \to \infty$$

Hilbert-Schmidt Operators

The operators K in the previous example is an example of a **Hilbert-Schmidt** operator (special class of compact operators).

Definition 3.31

Given a linear transformation $T : \mathcal{H} \to \mathcal{H}$ of the form

$$T(f)(x) = \int_{\mathbb{R}^n} K(x,y)f(y)\mathrm{d}y$$

is said to be **Hilbert-Schmidt** (HS) provided that $\|K\|_{\mathcal{L}^2(\mathbb{R}^n\times\mathbb{R}^n)} < \infty$.

Proposition 3.32

Let $T \in \mathrm{HS}(\mathbb{R}^n)$ with kernel $K(x,y)$.

1. Given $f \in \mathcal{L}^2(\mathbb{R}^n)$, $y \mapsto K(x,y)f(y) \in \mathcal{L}^1(\mathbb{R}^n)$ for a.e. $x \in \mathbb{R}^n$

2. $T : \mathcal{L}^2 \to \mathcal{L}^2$ is bounded with $\|T\| \leq \|K\|_{\mathcal{L}^2(\mathbb{R}^n \times \mathbb{R}^n)}$

3. T^* has kernel $\overline{K(y,x)}$

Proof.

1. By Fubini theorem, for almost every $x \in \mathbb{R}^n$, $y \mapsto |K(x,y)|^2 \in \mathcal{L}^1$ since by assumption

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |K(x,y)f(y)| \mathrm{d}y \mathrm{d}x < \infty$$

Now

$$\int_{\mathbb{R}^n} |K(x,y)||f(y)| \mathrm{d}y \leq \left(\int |K(x,y)|^2 \mathrm{d}y \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^n} |f(y)|^2 \mathrm{d}y \right)^{\frac{1}{2}} \tag{3.47}$$

by Cauchy-Schwartz for a.e. $x \in \mathbb{R}^n$. Since by assumption $\int_{\mathbb{R}^n \times \mathbb{R}^n} |K(x,y)|^2 \mathrm{d}x\mathrm{d}y$ where $\mathrm{d}x\mathrm{d}y$ is the product measure on product space, the iterated integrals for every slice are finite, and by Fubini, for a.e. $x \in \mathbb{R}_n$,

$$\int_{\mathbb{R}^n} |K(x,y)|^2 \mathrm{d}y < \infty \tag{3.48}$$

and substitution of (3.48) in (3.47) gives that for almost every $x \in \mathbb{R}^n$, $y \mapsto K(x,y)f(y) \in \mathcal{L}^1(\mathbb{R}^n_y)$.

2. Using Cauchy-Schwartz and applying (1) and Fubini,

$$\begin{aligned} \|Tf\|^2_{\mathcal{L}^2} &= \int \left(\int K(x,y)f(y)\mathrm{d}y \right) \overline{\left(\int K(x,y')f(y')\mathrm{d}y' \right)} \mathrm{d}x \\ &\leq \int \|K(x,\cdot)\|_{\mathcal{L}^2(y)} \cdot \|f\|_{\mathcal{L}^2(y)} \cdot \|f\|_{\mathcal{L}^2(y')} \mathrm{d}x \\ &= \|f\|^2_{\mathcal{L}^2} \int \|K(\cdot,\cdot)\|^2_{\mathcal{L}^2(\mathbb{R}^n \times \mathbb{R}^n)} \end{aligned}$$

and since the last part is

$$\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |K(x,y)|^2 \mathrm{d}y \right) \mathrm{d}x$$

and by Fubini, the iterated integrals equal the integral on the product space, so we have $\|K\|^2_{\mathcal{L}^2(\mathbb{R}^n \times \mathbb{R}^n)} \|f\|_{\mathcal{L}^2(\mathbb{R}^n)}$ therefore taking square roots on all terms

$$\frac{\|Tf\|_{\mathcal{L}^2}}{\|f\|_{\mathcal{L}^2}} \leq \|K\|_{\mathcal{L}^2(\mathbb{R}^n \times \mathbb{R}^n)}$$

3. Write $\langle Tf, g\rangle$ as a double integral and interchange orders of integration(by Fbini),

$$T^*f(x) = \int_{\mathbb{R}^n} \overline{K(x,y)} f(y)\mathrm{d}y$$

The proof is left as an exercise.

■

Example 3.8

Let $T : \mathcal{L}^2([-\pi, \pi]) \to \mathcal{L}^2([-\pi, \pi])$ with kernel

$$K(x,y) = \sum_{n\in\mathbb{Z}} \frac{e^{in(x-y)}}{1+n^2} = \sum_{n\in\mathbb{Z}} \frac{\cos(n(x-y))}{1+n^2}.$$

This is Hilbert-Schmidt.

Note

By Parseval, $\|K\|^2_{\mathcal{L}^2} = \sum_{n\in\mathbb{Z}} \left(\frac{1}{1+n^2}\right)^2 < \infty$. This implies that $T : \mathcal{L}^2([-\pi, \pi]) \Rightarrow \mathcal{L}^2([-\pi, \pi])$ is compact and HS.

Hilbert-Schmidt operators are special cases of what are called **compact** operators

Compact operators

Remark

Given $B = \{f \in \mathcal{H} : \|f\| \le 1\}$, the unit ball, this set is compact if $\dim \mathcal{H} < \infty$. However, if $\dim \mathcal{H} = \infty$, this is always false and B is **never** compact. B is compact thus if and only if $\dim \mathcal{H} < \infty$.

Definition 3.33 (Compactness)

We say that $X \subseteq \mathcal{H}$ is **compact** if for any sequence $\{f_n\} \subset X$, there exists a subsequence $\{f_{n_k}\} \subset \{f_n\}$ such that $\|f_{n_k} - f\| \to 0$ as $k \to \infty$ for some $f \in X$.

When $\dim \mathcal{H} = \infty$, consider the sequence $\{f_n\}_{n=-\infty}^{\infty} = \{e_n\}_{n=-\infty}^{\infty}$ where $e_n's$ are elements of the Hilbert basis, $\|e_n\| = 1$

Note

$\|f_n - f_m\| = \sqrt{2}$ if $n \neq m$ by Pythagoras, so there cannot exist a Cauchy subsequence.

Definition 3.34 (Compact operator)

$T : \mathcal{H} \to \mathcal{H}$ is compact provided that the closure of $T(B) \subset \mathcal{H}$,[10] $\mathrm{cl}(T(B))$ is compact,

[10] As $T(B)$ is precompact, its closure is then compact.

where

$$T(B) = \{g \in \mathcal{H} : g = Tf \text{ for } f \in B\}$$
$$B = \{f \in \mathcal{H} : \|f\| \leq 1\}$$

Proposition 3.35

Let $T : \mathcal{H} \to \mathcal{H}$ be compact. Then the following is true.

1. If $S : \mathcal{H} \to \mathcal{H}$ is compact, $ST : \mathcal{H} \to \mathcal{H}$ and $TS : \mathcal{H} \to \mathcal{H}$ are both compact. A two sided ideal in the space of operators (Fredholm operator)
2. Suppose $\{T_n\}_{n=1}^{\infty}$ are compact with $\|T_n - T\| \to 0$ as $n \to \infty$. Then $T : \mathcal{H} \to \mathcal{H}$ is compact.
3. Given $T : \mathcal{H} \to \mathcal{H}$ compact, there exists a finite rank operator $\{T_n\}_{n=1}^{\infty}$ with $\|T_n - T\| \to 0$ as $n \to \infty$
4. T is compact if and only if the adjoint T^* is compact.

Definition 3.36 (Finite-rank)

Suppose $\{e_n\}_{n\in\mathbb{Z}}$ is a Hilbert basis and $Te_n = \sum_m a_{mn}e_m$ and $S_n = \{a_{mn} \neq 0\}$ in $\mathcal{L}^2$. We say that $T : \mathcal{H} \to \mathcal{H}$ is finite rank provided there exists S with $\#S < \infty$ such that $S \supset \left(\bigcup_{n\in\mathbb{Z}} S_n\right)$. In other words, $(T) < \infty$ if for every $e_n \in \mathcal{H}$ (basis vector) $Te_n = \sum_{m\in\mathbb{Z}} a_{nm}e_m$ where $a_{nm} = 0$ except for finitely many m's (written as a finite-dimensional matrix). In addition, we require $\langle Te_n, Te_{n'}\rangle$ to be a a finite-dimensional matrix.

Here is some motivation: for smooth domain, the Dirichlet eigenvalue problem for smooth domain (can you hear the shape of a drum?) with $-\Delta\psi = \lambda^2\psi$ in some smooth domain Ω with $\psi|_{\partial\Omega} = 0$ boundary condition. For potential theory (or Fredholm theory), we may answer the question. The Dirichlet problem is still open; there are counterexample for polygonal domains, and some results in cases of symmetry. We would like to prove the following

Proof.

1. This is easy. Consider $TS : \mathcal{H} \to \mathcal{H}$. Suppose $\{f_n\}_{n=1}^{\infty} \in \mathcal{H}$ with $\|f_n\| \leq 1$. Since $S \in \text{com}(\mathcal{H})$, there exists $\{f_{n_k}\} \subset \{f_n\}$ such that $\|Sf_{n_k} = g\| \to 0$ as $k \to \infty$. But T is bounded,
$$\|T(Sf_{n_k}) - Tg\| \leq \|T\|\|Sf_{n_k} - g\| \to 0$$
as $k \to \infty$. Thus $TS \in \text{com}(\mathcal{H})$. To show that $ST \in \text{com}(\mathcal{H})$ is the same in essence and is left as an exercise.

2. $\{T_n\}_{n=1}^{\infty} \in \text{com}(\mathcal{H})$ with $\|T_n - T\| \to 0$ as $n \to \infty$. Given $\{f_n\} \in \mathcal{H}$ with $\|f_n\| < 1$, we first extract a convergent subsequence, using the sequence of compact operators $\{T_n\}_{n=1}^{\infty}$. **Diagonalization:** given f_n, since $T_1 \in \text{com}(\mathcal{H})$, by definition there a subsequence $\{f_{1,n}\} \subset \{f_n\}$ such that $T_1(f_{1,n})$ converges. Now, since T_2 is compact, $T_2 \in \text{com}(\mathcal{H})$, there exists a subsequence $\{f_{2,n}\} \subset \{f_1, n\}$ with $T_2(f_{2,n})$ convergent. Continue the process and let $g_k = f_{k,k}$ for $k = 1, 2, \ldots$.

   To visualize the process, consider the diagonal of the matrix consisting of the subsequences

$$\begin{pmatrix} \mathbf{f_{1,1}} & f_{1,2} & \cdots & \vdots \\ f_{2,1} & \mathbf{f_{2,2}} & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \mathbf{f_{k,k}} & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}$$

   Claim

   $\{Tg_k\}_{k=1}^{\infty}$ is Cauchy.

   Proof. This is an $\varepsilon/3$ argument:

$$\|Tg_k - Tg_l\| \le \|Tg_k - T_m g_k\| + \|T_m g_k - T_m g_l\| + \|T_m g_l - Tg_l\|$$

   using the triangle inequality. For any k, l

$$\|(T - T_m)g_k\| \to 0 \text{ as } m \to \infty \text{ and similarly } \|(T - T_m)g_l\| \to 0$$

   since by assumption $\|T - T_m\| \to 0$ as $n \to \infty$. For the second term, since T_m are bounded, for any fixed m

$$\|T_m g_k - T_m g_l\| \to 0 \text{ as } k, l \to \infty$$

   This implies that T is compact. ■

   Remark

   The point of the diagonalization is to ensure that for **any** $m \ge 1$, $\{T_m g_k\}_{k=1}^{\infty}$ is Cauchy. We need this to control to estimate the second term.

3. This is the finite-rank approximation: given $T \in \text{com}(\mathcal{H})$, we want to find finite rank $\{T_n\}_{n=1}^{\infty}$ with $\|T_n - T\| \to 0$ as $n \to \infty$. Let $\{e_k\}_{k=1}^{\infty} \in \mathcal{H}$ be a Hilbert basis. Let

$\Pi_n(f) = \sum_{k=1}^{n} \langle f, e_k \rangle e_k$. Π_n is an orthogonal projection on $\mathrm{span}\{e_k\}_{k=1}^{n}$. Now

$$Q_n = \mathrm{I_d} - \Pi_n$$
$$Q_n g = \sum_{k>n} a_k e_k \text{ where } g \overset{\mathcal{L}^2}{=} \sum_{k=1}^{\infty} a_k e_k \in \mathcal{H}$$

By Parseval,
$$\|Q_n g\| = \sum_{k>n} |a_k|^2.$$

Since $\|g\|^2 = \sum_{k=1}^{\infty} |a_k|^2 < \infty$, this implies the sequence of numbers $\{\|Q_n g\|^2\}_{n=1}^{\infty}$ decrease to zero as $n \to \infty$.

Idea: the candidate finite rank approximation to T is $\Pi_n T$ *i.e.* we have to show that $\|\Pi_n T - T\| \to 0$ as $n \to \infty$. Since $Q_n = \mathrm{I_d} - \Pi_n$, this is equivalent to showing that $\|Q_n T\| \to 0$ as $n \to \infty$.

■

Spectral theorem for compact operators

We have the following result, infinite dimension analog of the well-known linear algebra result

Theorem 3.37 (Spectral theorem)

Let $T : \mathcal{H} \to \mathcal{H}$ be a compact operator that is symmetric (*i.e.* $T^* = T$). Then, there exists a Hilbert basis $\{\varphi_k\}_{k=1}^{\infty}$ of $\mathcal{H}$ consisting of eigenvectors of T (*i.e.* $T\varphi_k = \lambda_k \varphi_k$). Moreover, the eigenvalues λ_k are real and $\lambda \to 0$ as $k \to \infty$.

Before proving this, we motivate the result

Example 3.9

Let

$$K(x, y) = \sum_{n \in \mathbb{Z}} \frac{e^{in(x-y)}}{n^2 + 1} \quad \text{and} \quad Tf(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} K(x, y) f(y) \mathrm{d}y,$$

an Hilbert-Schmidt operator and $T : \mathcal{L}^2([-\pi, \pi]) \to \mathcal{L}^2([-\pi, \pi])$. The eigenvectors are $\{e^{inx}\}_{n \in \mathbb{Z}}$ and the eigenvalues $\left\{\frac{1}{n^2+1}\right\}_{n \in \mathbb{Z}}$. Here $\|T\| = 1$ is also an eigenvalue (this is not a mere coincidence), along with $\frac{1}{2}, \frac{1}{5}$, etc.

In addition, we will show that if $T \in \mathrm{com}(\mathcal{H})$ and symmetric, then either $\|T\|$ or $-\|T\| \in \mathrm{spec}(T)$, where spec, or spectrum, is the set of eigenvalues. In particular, $\mathrm{spec}(T) \neq \emptyset$.

Note

If $T \in \text{com}(\mathcal{H})$, but T is not symmetric, these results are false.

Example 3.10 (Volterra operators)

In the simple case, we have $T : \mathcal{L}^2([0,1]) \to \mathcal{L}^2([0,1])$ with $Tf(x) = \int_0^x f(y)\mathrm{d}y$ for $x \in [0,1]$. We will show (in assignment) that

1. T is compact ($T \in \text{com}(\mathcal{L}^2([0,1])$ (easy as it is Hilbert-Schmidt)
2. $\text{spec}(T) = \{0\}$. Thus, there are no non-trivial eigenvalues of this operator. In particular, $T^* \neq T$.

Proof. The first step consists of the following lemma

Lemma 3.38

Let $T : \mathcal{H} \to \mathcal{H}$ be bounded and symmetric ($T^* = T$).

1. If λ is an eigenvalue of T, then $\lambda \in \mathbb{R}$
2. Let f_1, f_2 be eigenvectors corresponding to distinct eigenvalues $\lambda_1 \neq \lambda_2$. Then $\langle f_1, f_2 \rangle = 0$.

Proof.

1. Let φ be a non-trivial eigenvector with $T\varphi = \lambda\varphi$ with $\lambda \neq 0$. Then

$$\begin{aligned}
\lambda \langle \varphi, \varphi \rangle &= \langle \lambda\varphi, \varphi \rangle && \text{(by linearity)} \\
&= \langle \varphi, T\varphi \rangle && \text{(since } T^* = T\text{)} \\
&= \overline{\lambda} \langle \varphi, \varphi \rangle
\end{aligned}$$

which implies that $\lambda = \overline{\lambda}$ so that $\lambda \in \mathbb{R}$.

2. Suppose $T\varphi_1 = \lambda_1\varphi_1$ and $T\varphi_2 = \lambda_2\varphi_2$, then

$$\lambda_1 \langle \varphi_1, \varphi_2 \rangle = \langle \varphi_1, T\varphi_2 \rangle$$

and since $\lambda_1 \neq \lambda_2$, this implies $\langle \varphi_1, \varphi_2 \rangle = 0$ by part 1, since $\lambda_1, \lambda_2 \in \mathbb{R}$.

■

The second step consists in characterization of the eigenspaces.

Lemma 3.39

Let $T \in \text{com}(\mathcal{H}), T^* = T$ and let $\lambda \neq 0$. Then

1. $\dim \ker(T - \lambda \mathrm{I_d}) < \infty$.

2. For any $\mu > 0$, $\dim_{\lambda_k > \mu} V_{\lambda_k} < \infty$. Here, V_{λ_k} is the vector space generated by eigenfunctions with eigenvalue λ_k.

3. $\mathrm{spec}(T) = \{\lambda_k\}_{k=1}^{\infty}$ where $\lambda_k \to 0$ as $k \to \infty$.

Proof.

1. Let $V_\lambda = \ker(T - \lambda \mathrm{I_d})$. Suppose that $\dim V_\lambda = \infty$. Then, there exists an orthonormal set $\{\varphi_k\}_{k=1}^{\infty} \in V_\lambda$ with $T\varphi_k = \lambda \varphi_k$ for $k = 1, 2, \ldots$. By rescaling, let $\|\varphi_k\| = 1$. Since $T \in \mathrm{com}(\mathcal{H})$, there is a subsequence $\{\varphi_{n_k}\}$ such that $T\varphi_{n_k} \to g$ as $n \to \infty$. Then

$$\|T\varphi_{n_k} - T\varphi_{n_l}\| = \lambda \|\varphi_{n_k} - \varphi_{n_l}\| = \sqrt{2}\lambda$$

   if $k \neq l$ and $\lambda \neq 0$. Thus, $\{T\varphi_{n_k}\}$ cannot converge. Contradiction.

2. Similar to 1. We argue by contradiction; fix $\mu > 0$ and consider V_{λ_k} where $\lambda_k > \mu$. Since eigenvalues are distinct (by Lemma 3.38), we choose an orthogonal (wlog orthonormal) set of vectors $\{\varphi_n\}_{n=1}^{\infty}$ spanning $\bigoplus_{\lambda_k > \mu} V_{\lambda_k}$. Since T is compact, there exists $\{\varphi_{n_k}\} \Rightarrow T\varphi_{n_k}$ converge as $k \to \infty$. Then $T\varphi_{n_k} = \lambda_{n_k}\varphi_{n_k}$ with $\lambda_{n_k} > \mu$. Then

$$\|T\varphi_{n_k} - T\varphi_{n_l}\|^2 = \|\lambda_{n_k}\varphi_{n_k} - \lambda_{n_l}\varphi_{n_l}\|^2 = \lambda_{n_k}^2 + \lambda_{n_l}^2 > 2\mu^2 > 0$$

   and this is a contradiction, implying that $\dim \left(\bigoplus_{\lambda_k > \mu} V_{\lambda_k} \right) < \infty$

3. There are 2 points outstanding, namely

   - $\mathrm{spec}(T) \neq \emptyset$ for $T \neq 0$
   - $\bigoplus_{k=1}^{\infty} V_{\lambda_k} = \mathcal{H}$ with $V_{\lambda_k} = \{\phi_k \in \mathcal{H} \mid T\phi_k = \lambda_k \phi_k\}$

■

We begin with the first claim

Lemma 3.40
Assume $T \in \mathrm{com}(\mathcal{H})$ and $T \neq 0$ symmetric ($T = T^*$). Then $\mathrm{spec}(T) \cap \{\pm\|T\|\} \neq \emptyset$. Either $\|T\|$ or $-\|T\| \in \mathrm{spec}(T)$.

Proof. Using the polarization identity, one can show (exercise) when $T = T^*$ that

$$\|T\| = \sup \{ | \langle Tf, f \rangle | ; \|f\| = 1 \} \tag{3.49}$$

Recall, for general operators,

$$\|T\| = \sup\{|\langle Tf, g\rangle|; \|f\| \leq 1, \|g\| \leq 1\}$$

see the polarization identity in the book. Symmetry is crucial here. From (3.49), either $\|T\| = \sup\{\langle Tf, f\rangle\,;\|f\| = 1\}$ or $-\|T\| = \inf\{\langle Tf, f\rangle\,;\|f\| = 1\}$. Wlog, we assume $\|T\| = \sup\{\langle Tf, f\rangle\,;\|f\| = 1\}$.

So we can find a sequence of vectors $\{f_n\} \in \mathcal{H}$ with $\|f\| = 1$ such that $\langle Tf_n, f_n\rangle \xrightarrow{n\to\infty} \lambda$ as $\lambda = \|T\|$. Since $T \in \text{com}(\mathcal{H})$, by passing to a subsequence, we have that $\|Tf_n - g\| \to 0$ as $n \to \infty$ where $g \in \mathcal{H}'$.

Claim

$g \neq 0$ is an eigenvector of T with $Tg = \lambda g$. Using symmetry, if we look at

$$\begin{aligned}\|Tf_n - \lambda f_n\|^2 &= \|Tf_n\|^2 + \lambda^2\|f\|^2 - 2\lambda\langle Tf_n, f_n\rangle \\ &\leq \|T\|^2\|f_n\|^2 + \lambda^2\|f_n\|^2 - 2\lambda\langle Tf_n, f_n\rangle \\ &= 2\lambda^2\|f_n\|^2 - 2\lambda\langle Tf_n, f_n\rangle \\ &= 2\lambda^2 - 2\lambda\langle Tf_n, f_n\rangle \\ &\to 2\lambda^2 - 2\lambda^2\end{aligned}$$

since $\langle Tf_n, f_n\rangle \to \lambda$. So the upshot is that

$$\|Tf_n - \lambda f_n\| \to 0 \quad \text{as } n \to \infty. \tag{3.50}$$

Since $\|Tf_n - g\| \to 0$ as $n \to \infty$ from (3.50),

$$\|g - \lambda f_n\| \leq \|Tf_n - \lambda f_n\| + \|Tf_n - g\| \to 0.$$

As $n \to \infty$,

$$\begin{aligned}\|Tg - \lambda Tf_n\| &\leq \|T\|\|g - \lambda f_n\| \to 0 \\ \|\lambda g - \lambda Tf_n\| &\leq \lambda\|g - Tf_n\| \quad\ \to 0\end{aligned}$$

which implies $Tg = \lambda g$ if $g \neq 0$.

■

Let $S = \overline{\bigoplus_{k=1}^{\infty} V_{\lambda_k}} \neq \emptyset$ by the previous lemma (Lemma 3.40). By the decomposition theorem for $\mathcal{H}$, if $\mathcal{H} = S$, we are done. Assume that $\mathcal{H} \neq S$. Then, there exists $\{0\} \neq S^\perp \subset \mathcal{H}$ closed with $\mathcal{H} = S \oplus S^\perp$. We want to show that there is no $S^\perp$. The key point here is that T preserves the decomposition $S \oplus S^\perp$ and $T : S \to S$ and $T : S^\perp \to S^\perp$ and if we have an eigenspace, then T preserves the eigenspace.

If $T(T\varphi_k) = T(\lambda_k \varphi_k) = \lambda_k T\varphi_k$ implying that $T : V_{\lambda_k} \to V_{\lambda_k}$. For the second statement, suppose $g \in S^\perp$ and $T\varphi_k = \lambda_k \varphi_k$. Then

$$0 = \langle g, T\varphi_k \rangle = \langle Tg, \varphi_k \rangle \text{ using } T^* = T;$$

then $T : S^\perp \to S^\perp$. Consider $T\big|_{S^\perp} : S^\perp \to S^\perp$. Clearly, $(T\big|_{S^\perp})^* = (T\big|_{S^\perp})$ and $T\big|_{S^\perp}$ is compact. Then, by the argument we have so far, there exists $\lambda' \neq 0 \in \mathbb{R}$ with $T\big|_{S^\perp}\varphi = \lambda'\varphi$, with $\varphi \in S^\perp \neq 0$, implying $T\varphi = \lambda'\varphi$. Contradiction. ■

Let us look at some examples of compact operators that are not Hilbert Schmidt

For example, the singular integral convolution operators of the form $T : \mathcal{L}^2([0,1]) \to \mathcal{L}^2([0,1])$, for *e.g.* $Tf(x) = \int_0^1 |x-y|^{-\frac{1}{2}} f(y)\mathrm{d}y$.

Claim
$T \in \mathrm{com}(\mathcal{L}^2([0,1]))$.

Note
$K(x,y) = |x-y|^{-\frac{1}{2}}$ and $K \notin \mathcal{L}^2([0,1] \times [0,1])$ implies that T is not Hilbert-Schmidt.

Proof. We use the result that says that $\|T_n - T\| \to 0$ as $n \to \infty$ with $T_n \in \mathrm{com}(\mathcal{H})$ implies that $T \in \mathrm{com}(\mathcal{H})$. Consider here the family of operators

$$T_\varepsilon f(x) = \int_0^1 (|x+y| + \varepsilon)^{-\frac{1}{2}} f(y)\mathrm{d}y$$

for $\varepsilon > 0, \varepsilon \in \{n^{-1}\}_{n=1}^\infty$. For any $\varepsilon > 0$, $K_\varepsilon(x,y) = (|x-y| + \varepsilon)^{-\frac{1}{2}} \in \mathcal{C}^0([0,1] \times [0,1])$, therefore T_ε is Hilbert-Schmidt and $T_\varepsilon \in \mathrm{com}(\mathcal{L}^2([0,1]))$. We need to prove $\|T_\varepsilon - T\| \to 0$ as $\varepsilon \to 0$.

$$\begin{aligned}
\|(T_\varepsilon - T)f\|_{\mathcal{L}^2}^2 &= \int_0^1 |(T_\varepsilon - T)f(x)|^2 \mathrm{d}x \\
&\leq \int_0^1 \left| \int_0^1 \big(K_\varepsilon(x,y) - K(x,y)\big) f(y)\mathrm{d}y \right|^2 \mathrm{d}x \\
&\leq \int_0^1 \left(\int_0^1 |K_\varepsilon(x,y) - K(x,y)|^2 \mathrm{d}y \right) \|f\|_{\mathcal{L}^2}^2 \mathrm{d}x \\
&= \|f\|_{\mathcal{L}^2}^2 \int_0^1 \int_0^1 |K_\varepsilon(x,y) - K(x,y)|^2 \mathrm{d}y\mathrm{d}x
\end{aligned}$$

using Fubini and Cauchy-Schwartz. Let $z = x - y$; we are reduced to estimating

$$\int_{-1}^{1} |K_\varepsilon(z) - K(z)|^2 \mathrm{d}z = \int_{-1}^{1} \left|(z+\varepsilon)^{-\frac{1}{2}} - z^{-\frac{1}{2}}\right|^2 \mathrm{d}z$$
$$= \int_{|z|<10\varepsilon} \left|(z+\varepsilon)^{-\frac{1}{2}} - z^{-\frac{1}{2}}\right|^2 \mathrm{d}z + \int_{|z|\geq 10\varepsilon}^{1} \left|(z+\varepsilon)^{-\frac{1}{2}} - z^{-\frac{1}{2}}\right|^2 \mathrm{d}z$$

Now for $|z| > 10\varepsilon$,

$$(z+\varepsilon)^{-\frac{1}{2}} = z^{-\frac{1}{2}} \left(1 + \frac{\varepsilon}{z}\right)^{-\frac{1}{2}}$$
$$= z^{-\frac{1}{2}} \left(1 - \frac{1}{2}\frac{\varepsilon}{z} + O\left(\left(\frac{\varepsilon}{z}\right)^2\right)\right)$$

and

$$\left|(z+\varepsilon)^{-\frac{1}{2}} - z^{-\frac{1}{2}}\right| = z^{-\frac{1}{2}} \left|\frac{1}{2}\frac{\varepsilon}{z^{\frac{3}{2}}} + O\left(\frac{\varepsilon}{z}\right)^2 z^{-\frac{1}{2}}\right|$$

Note

$$\int_{1>|z|>10\varepsilon} \left|\frac{1}{2}\frac{\varepsilon}{z} + O\left(\frac{\varepsilon}{z}\right)^2\right| \mathrm{d}z \to 0 \text{ as } \varepsilon \to 0.$$

Exercise 3.1

Using integrability property of $z \mapsto z$, one can show that the first term goes to zero as $\varepsilon \to 0^+$ as well.

■

These are called singular integral operators.

We are now going back to digress for the Dirichlet problem. We will show it for $\mathbb{R}^3$ for general bounded domain.

Example 3.11 (Dirichlet problem)

Let $D \subset \mathbb{R}^3$ be a bounded domain with smooth boundary, *i.e.* ∂D is $\mathcal{C}^\infty$. Recall the Dirichlet problem is of the form

$$\begin{cases} \Delta u = 0 \text{ in } D \\ u\big|_{\delta\Lambda} = f \in \mathcal{C}^\infty \end{cases}$$

Then, apply the **method of boundary layers**

We want to find a function $f(x, y)$ singular at $x = -y$

$$\int_D (\Delta_y u)(y) G(x,y) \mathrm{d}y - \int_D u(y) \Delta_y G(x,y) \mathrm{d}y$$
$$= \int_{\partial D} \partial_{\nu_y} u(y) G(x,y) d\sigma(y) - \int_{\partial D} u(y) \partial_y G(x,y) d\sigma(y)$$

and $\langle \nu_y, G(x,y) \rangle = \partial_{\nu_y} G(x,y)$. The term $(\Delta_y u)(y) G(x,y)$ is zero. Also, $\partial_{\nu_y} u(y) G(x,y) = 0$; this is non-trivial, but follows from argument. One then gets an integral expression of the form

$$\int_D u(y) \Delta_y G(x,y) \mathrm{d}y = \int_{\partial D} \partial_{\nu_y} G(x,y) u_y \mathrm{d}\sigma(y)$$

We choose $G(x, y)$ to be the Greens function

$$G(x,y) = (4\pi)^{-1} |x - y|^{-1} \tag{3.51}$$

and G has the property that

$$\Delta_y G(x,y) = 0; x \neq y$$

and moreover,

$$\Delta_y G(x,y) = \delta(x - y) \tag{3.52}$$

where (3.52) means that

$$\int_D \Delta_y G(x,y) d(y) \mathrm{d}y = f(x)$$

With the choice of $G(x, y)$ in (3.51) for $x \in \mathbb{R}^d$

$$u(x) = \int_{\partial D} K(x,y) \underbrace{f(y)}_{u|_{\delta\Lambda}(y)} \mathrm{d}\sigma(y) \tag{3.53}$$

Since $K(x, y)$ is singular at $x = y$, one cannot take limits inside the integral (3.53) as $x \to x_0 \in \partial D$. One can show that one actually gets the following equation

$$-\varphi(x_0) + \int_{\partial D} K(x_0, y_0) f(y_0) \mathrm{d}\sigma(y_0) = f(x)$$

We want to find $\varphi \in \mathcal{C}^0(\partial D)$ solving $(-\mathrm{I_d} + T)\varphi = f$

$$T\varphi(x) = \int_{\partial D} K(x,y) f(y) \mathrm{d}\sigma(y) \quad \text{for} x \in \partial D$$

and $T : \mathcal{C}^0(\partial D) \to \mathcal{C}^0(\partial D)$ is a **compact operator.** There is nothing in the kernel, meaning that there is a unique solution to the problem.

Section 4
Fourier transforms

Let $f : \mathbb{R} \to \mathbb{C}$ be a function and the corresponding Fourier transform, denoted $\hat{f} : \mathbb{R} \to \mathbb{C}$ and defined as

$$\hat{f}(y) = \int f(x)e^{-ixy}\mathrm{d}x$$

Remark

If $f \in \mathcal{L}^1(\mathbb{R})$, then the Fourier transform is well defined as $|\hat{f}(y)| \leq \|f\|_{\mathcal{L}^1(\mathbb{R})}$. If $f \in \mathcal{L}^2(\mathbb{R})$, then this is not necessarily the case. For example, consider $f(x) = x^{-\frac{3}{4}} \notin \mathcal{L}^1(\mathbb{R})$ and the Fourier transform may not be well-defined; we don't know if the integral defining $\hat{f}$ is well-defined.

We will thus restrict ourselves to the (smaller) space of Schwartz functions,

$$\mathcal{S} = \left\{ f \in \mathcal{C}^\infty(\mathbb{R}) : \ \forall\, n, m \in \mathbb{Z}^+, \exists\, c_{n,m} \geq 0 \text{ such that } \left\| x^n \frac{\partial^m}{\partial x^m} f \right\|_\infty \leq c_{n,m} \right\}.$$

and we have the following properties (closure under additivity and linearity), namely if $f, g \in \mathcal{S}$, then $f + g \in \mathcal{S}$ and if $c \in \mathbb{R}$, then $f \in \mathcal{S}$. From the triangle inequality, we have for any polynomial $f \in \mathcal{S}$ that

$$\left\| p(x) \frac{\partial^m f}{\partial x^m} \right\| \leq c_m$$

for a polynomial $p(x)$.

Remark

If $f \in \mathcal{S}$, then for $\forall\, N, \exists c_N > 0$ such that

$$|f(x)| \leq \frac{c}{(1 + |x|^2)^N} \tag{4.54}$$

Remark

If $f \in \mathcal{S}$, then by (4.54) $f \in \mathcal{L}^1(\mathbb{R})$ and as such $\hat{f}$ is well-defined.

Now, if $f \in \mathcal{S}$, then $\hat{f} \in \mathcal{S}$. The map $\hat{\cdot} : \mathcal{S} \to \mathcal{S}$ is an isomorphism.

Lemma 4.1

1. If $f \in \mathcal{S}$, then $\frac{\mathrm{d}}{\mathrm{d}y}\hat{f}(y) = -i\widehat{xf(y)}$
2. If $f \in \mathcal{S}$, then $\widehat{\frac{\mathrm{d}f}{\mathrm{d}x}}(y) = iy\hat{f}(y)$

namely, it exchanges differentiation with products.

Proof. Recall that

$$\hat{f}(y) = \int f(x)e^{-ixy}\mathrm{d}x$$

1. e^{-ixy} is differentiable with respect to y, therefore $\hat{f}$ is differentiable too.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}y}\hat{f}(y) &= \int \frac{\mathrm{d}}{\mathrm{d}y}\left(f(x)e^{-ixy}\right)\mathrm{d}x \\ &= -i\int xf(x)e^{-ixy}\mathrm{d}x \\ &= -i\widehat{xf}(y)\end{aligned}$$

2. We have this time integrating by parts (since (4.54) holds)

$$\begin{aligned}\widehat{\frac{\mathrm{d}f}{\mathrm{d}x}}(y) &= \int \frac{\mathrm{d}f}{\mathrm{d}x}e^{-ixy}\mathrm{d}x \\ &= -\int f(x)\frac{\mathrm{d}}{\mathrm{d}x}e^{-ixy}\mathrm{d}x \\ &= iy\int f(x)e^{-ixy}\mathrm{d}x \\ &= iy\hat{f}(y)\end{aligned}$$

■

Lemma 4.2

If $f \in \mathcal{S}$, then $\hat{f} \in \mathcal{S}$. Both $y\hat{f}(y)$ and $\frac{\partial \hat{f}(y)}{\partial y}(y)$ are Fourier transforms of another function in $\mathcal{S}$ by the previous lemma. By induction, the function $y^n\frac{\partial^m}{\partial y^m}\hat{f}(y) = \hat{g}_{n,m}$ for some $g_{n,m} \in \mathcal{S}$. We have

$$|\hat{g}_{n,m}(y)| \le \|g_{n,m}\|_{\mathcal{L}^1(\mathbb{R})} \coloneqq c_{n,m} \quad \Rightarrow \left\|y^n\frac{\partial^m}{\partial x^m}\hat{f}(y)\right\| \le c_{n,m}$$

which implies that $\hat{f} \in \mathcal{S}$.

Example 4.1

Consider the Gaussian function $f(x) = e^{-\frac{x^2}{2}}$. The Fourier transform is given by $\hat{f}(y) = \sqrt{2\pi}e^{-\frac{y^2}{2}}$.

First, differentiate the function with respect to x:

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = -xf(x) \quad \Rightarrow \quad \widehat{\frac{\mathrm{d}f}{\mathrm{d}x}}(y) = -\widehat{xf}(y)$$

Then, by Lemma 4.1,

$$iy\hat{f}(y) = -i\frac{\mathrm{d}}{\mathrm{d}y}\hat{f}(y) \Rightarrow \frac{\mathrm{d}}{\mathrm{d}y}\hat{f}(y) = -y\hat{f}(y)$$

and both f and $\hat{f}$ satisfy the differential equation

$$\frac{\mathrm{d}u}{\mathrm{d}x} + xu = 0 \tag{4.55}$$

If u satisfies (4.55), then

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(e^{\frac{x^2}{2}}u(x)\right) = e^{\frac{x^2}{2}}\left(xu(x) + \frac{\mathrm{d}u}{\mathrm{d}x}(x)\right) = 0$$

which implies that $e^{\frac{x^2}{2}}u(x) = c$ for some $c \in \mathbb{R}$, where $u(x) = \hat{f}(x)$. Now $\hat{f}(y) = ce^{-\frac{y^2}{2}}$ and find c. We find that $c = \hat{f}(0) = \int f(x)e^{-ix0}\mathrm{d}x = \int e^{-\frac{x^2}{2}}\mathrm{d}x = \sqrt{2\pi}$. Thus, $\hat{f}(y) = \sqrt{2\pi}e^{-\frac{y^2}{2}}$.

Lemma 4.3

If $f, g \in \mathcal{S}$, then

$$\int \hat{f}(y)g(y)\mathrm{d}y = \int f(x)\hat{g}(x)\mathrm{d}x$$

Proof.

$$\begin{aligned}\int \hat{f}(y)g(y)\mathrm{d}y &= \int \left(\int f(x)e^{-ixy}\mathrm{d}x\right) g(y)\mathrm{d}y \\ &= \int\int f(x)g(y)e^{-ixy}\mathrm{d}x\mathrm{d}y \\ &= \int\int f(x)g(y)e^{-ixy}\mathrm{d}y\mathrm{d}x \\ &= \int \left(\int g(y)e^{-ixy}\mathrm{d}y\right) f(x)\mathrm{d}x \\ &= \int \hat{g}(x)f(x)\mathrm{d}x\end{aligned}$$

using (4.54). ■

Lemma 4.4

If $f \in \mathcal{S}$ and $a \in \mathbb{R}$, then $f_a(x) = f(x+a)$. Then $\hat{f}_a = e^{iay}\hat{f}$.

Proof.

$$\begin{aligned}\hat{f}_a(y) &= \int f_a(x)e^{-ixy}\mathrm{d}x \\ &= \int f(x+a)e^{-ixy}\mathrm{d}x\end{aligned}$$

Make the change of variable $s = x + a$

$$= \int f(s)e^{-(s-a)y}\mathrm{d}s = e^{iay}\int f(x)e^{-isy}\mathrm{d}s = e^{iay}\hat{f}(y)$$

■

Lemma 4.5

If $f \in \mathcal{S}$ and $a > 0$, then $f_a(x) = f\left(\frac{x}{a}\right)$ implies $\hat{f}_a(y) = a\hat{f}(ay)$.

Proof. By making the change of variable $s = \frac{x}{a}$, we have

$$\begin{aligned}\hat{f}_a(y) &= \int f\left(\frac{x}{a}\right)e^{ixy}\mathrm{d}x \\ &= \int f(s)e^{-isay}a\mathrm{d}s \\ &= a\int f(s)e^{-say}\mathrm{d}s \\ &= a\hat{f}(ay)\end{aligned}$$

■

The Fourier transform map is a bijection and is linear. The inverse map

$$f(x) = \frac{1}{2\pi}\int \hat{f}(y)e^{ixy}\mathrm{d}y = \frac{1}{2\pi}\int g(y)e^{ixy}\mathrm{d}y$$

the so-called inverse Fourier transform of g, denoted $g^{\vee}(y)$ and $(\hat{f})^{\vee} = f$.

4.1 Fourier Transform

We now tackle the proof of the inverse Fourier transforms, in $\mathbb{R}$ for simplicity. There are parallels and differences between the continuous and discrete Fourier transforms. The

Schwartz functions are

$$\mathcal{S}(\mathbb{R}) = \left\{ f \in C^{\infty}(\mathbb{R}) : \left| x^m \left(\frac{\mathrm{d}}{\mathrm{d}x} \right)^n f(x) \right| \leq c_{n,m} < \infty \right\} \text{ for all } m, n \geq 0$$

Example 4.2

$C_0^{\infty}(\mathbb{R}) \subset \mathcal{S}(\mathbb{R})$, the compactly supported smooth functions are smooth functions. Glue together piecewise functions and $e^{-\frac{1}{x^2}}$ functions, which are smooth.

Example 4.3

The Gaussian function $f(x) = e^{-\frac{x^2}{2}}$ and for any polynomial decay at infinitum.

Definition 4.6

The Fourier transform $f : \mathcal{S} \to \mathcal{S}$ where $\mathcal{S} = \mathcal{S}(\mathbb{R})$ is written

$$(\mathcal{F}g)(y) = \int_{\mathbb{R}} e^{-ixy} g(x) dx$$

We sometimes write $(\mathcal{F}g)(y) = \hat{g}(y)$. It is easy to check by dominated convergence and the fact that $g \in \mathcal{S}(\mathbb{R})$ that $\mathcal{F} : \mathcal{S} \to \mathcal{S}$.

This follows from some basic facts about Fourier transforms Given $f \in \mathcal{S}$, if $g(x) = xf(x)$, then $\hat{g}(y) = \sqrt{-1}\frac{\mathrm{d}}{\mathrm{d}y}\hat{f}(y)$ and conversely, given $f \in \mathcal{S}$ and $h(x) = \frac{\mathrm{d}}{\mathrm{d}x} f(x)$, $\hat{h}(y) = \sqrt{-1} g\hat{f}(y)$. and $\hat{f}(y) = \int e^{-ixy} f(x) \mathrm{d}x$.

Theorem 4.7 (Fourier inversion)

$\mathcal{F} : \mathcal{S} \to \mathcal{S}$ is bijective onto $\mathcal{S}$, and moreover, if $f \in \mathcal{S}$ and $g = \hat{f}$, then $f = \check{g}$ where

$$\check{g}(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ixy} g(y) \mathrm{d}y$$

and $(\hat{f}\,)^{\vee} = f$.

Remark

We note

$$\check{g}(y) = \frac{1}{2\pi} \hat{g}(-y)$$

for $g \in \mathcal{S}$.

Lemma 4.8

Given $f, g \in \mathcal{S}$, then

$$\int_R \hat{f}(y) g(y) \mathrm{d}y = \int_{\mathbb{R}} f(x) \hat{g}(x) \mathrm{d}x$$

Proof. We make an application of Fubini

$$\begin{aligned}\int_{\mathbb{R}} \hat{f}(y)g(y)\mathrm{d}y &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-ixy} f(x)\mathrm{d}x\right) g(y)\mathrm{d}y \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-ixy} g(y)\mathrm{d}y\right) f(x)\mathrm{d}x \\ &= \int_{\mathbb{R}} \hat{g}(x)f(x)\mathrm{d}x\end{aligned}$$

Note

$F(x, y) = e^{-ixy} f(x)g(y) \in \mathcal{L}^1(\mathbb{R}^n \times \mathbb{R}^n)$

∎

The other 2 elementary properties of $\mathcal{F}$ that we use in the proof of the Fourier inversion theorem have to do with how the Fourier transforms behaved relative to translation and scaling.

If $g(y) = e^{-\frac{y^2}{2a^2}}$, then

$$\hat{g}(x) = \sqrt{2\pi} a e^{-\frac{a^2 x^2}{2}} \tag{4.56}$$

the explicit formula. Using (4.56) in Lemma 4.8 with $g(y) = e^{-\frac{y^2}{2a^2}}$, then

$$\int_{\mathbb{R}} \hat{f}(y) e^{-\frac{y^2}{2a^2}} \mathrm{d}y = \sqrt{2\pi} a \int_{\mathbb{R}} f(x) e^{-\frac{a^2 x^2}{2}} \mathrm{d}x$$

The idea is to take $a \to \infty$ in both sides of (4.1) after making a change of variable $x \mapsto ax \coloneqq s$ on the right hand side, which becomes

$$\sqrt{2\pi} \int_{\mathbb{R}} f\left(\frac{s}{a}\right) e^{-\frac{s^2}{2}} \mathrm{d}s.$$

The upshot is that

$$\int_{\mathbb{R}} \hat{g}(y) e^{-\frac{y^2}{2a^2}} \mathrm{d}y = \sqrt{2\pi} \int_{\mathbb{R}} f\left(\frac{s}{a}\right) e^{-\frac{s^2}{2}} \mathrm{d}s \tag{4.57}$$

we take $a \to \infty$ limit of both sides. Since $f \in \mathcal{S}$, then $\hat{f} \in \mathcal{S}$ and by dominated convergence theorem, the LHS is

$$\lim_{a\to\infty}\left(\int_{\mathbb{R}} \hat{g}(y)e^{-\frac{y^2}{2a^2}}\mathrm{d}y\right) = \int_{\mathbb{R}} \hat{f}(y)\mathrm{d}y$$

As for the right hand side, we again apply dominated convergence theorem to get

$$\lim_{a\to\infty}\left(\sqrt{2\pi}\int_{\mathbb{R}} f\left(\frac{s}{a}\right)e^{-\frac{s^2}{2}}\mathrm{d}s\right) = \sqrt{2\pi}\left(\int_{\mathbb{R}} e^{-\frac{s^2}{2}}\mathrm{d}s\right)f(0) = 2\pi f(0).$$

What we have so far is a special case of the inversion formula; we have proved that for any $f \in \mathcal{S}$,

$$f(0) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}(x)\mathrm{d}x.$$

Now, replace $f(x)$ by $h(x) = f(x+a)$ for $a \in \mathbb{R}$. Then

$$h(0) = f(a) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{h}(x)\mathrm{d}x = \frac{1}{2\pi}\int_{\mathbb{R}} e^{iax}\hat{f}(x)\mathrm{d}x$$

for $a \in \mathbb{R}$.

Theorem 4.9 (Plancherel)

Given $f \in \mathcal{S}(\mathbb{R})$, then

$$\|\hat{f}\|^2_{\mathcal{L}^2(\mathbb{R})} = 2\pi\|f\|^2_{\mathcal{L}^2(\mathbb{R})}$$

Notation

We use the notation $\hat{f}$ and $\mathcal{F}(f)$ both for Fourier transform. Similarly, the inverse transform

$$\check{f}(x) = \mathcal{F}^{-1}(f)(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{ixy}\hat{f}(y)\mathrm{d}y$$

The inversion formula simply says that

$$f = \mathcal{F}^{-1}(\mathcal{F}(f)), \qquad \text{for } f \in \mathcal{S}(\mathbb{R})$$

Proof. We use the identity

$$\int_{\mathbb{R}} f\hat{g}\mathrm{d}x = \int_{\mathbb{R}} \hat{f}g\mathrm{d}x \quad \text{for } f, g, \in \mathcal{S}(\mathbb{R}) \tag{4.58}$$

and we choose $g(x)$ in a judicious way. By the inversion formula, for $f \in \mathcal{S}$

$$\begin{aligned} f(x) &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{ixy}\hat{f}(y)\mathrm{d}y \\ \Rightarrow \overline{f(x)} &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{-ixy}\overline{\hat{f}(y)}\mathrm{d}y = \frac{1}{2\pi}\mathcal{F}(\overline{\hat{f}}) = \frac{1}{2\pi}\mathcal{F}\left(\overline{\mathcal{F}(f)}\right) \end{aligned}$$

which implies that

$$\mathcal{F}(\overline{\mathcal{F}(f)}) = 2\pi\overline{f} \tag{4.59}$$

Let $f = \overline{\mathcal{F}(f)} = \overline{\hat{f}}$. By (4.58), the RHS is

$$= \int_{\mathbb{R}} \hat{f}g = \int_{\mathbb{R}} \hat{f}\overline{\hat{f}} = \int_{\mathbb{R}} |\hat{f}|^2$$

and the LHS is simply

$$\begin{aligned} &= \int_{\mathbb{R}} f\hat{g}\mathrm{d}x \\ &= \int_{\mathbb{R}} f\mathcal{F}(\overline{\mathcal{F}(f)})\mathrm{d}x \\ &= \int_{\mathbb{R}} f\left(\overline{\hat{f}}\right)^{\wedge}\mathrm{d}x \\ &= 2\pi\int_{\mathbb{R}} f\overline{f}\mathrm{d}x \\ &= 2\pi\int_{\mathbb{R}} |f|^2 \end{aligned}$$

by (4.59). We can use Plancherel formula to compute the norm when either side of $2\pi\int|f|^2 = \int|\hat{f}|^2$ is easier than the other to evaluate. ■

4.2 Extension of Fourier transform to $\mathcal{L}^2(\mathbb{R})$

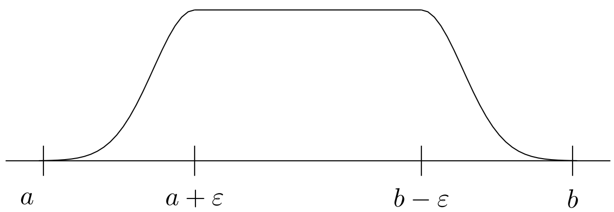
Remark

Given $f \in \mathcal{L}^2(\mathbb{R})$,

$$\hat{f}(y) = \int_{\mathbb{R}} e^{-ixy}f(x)\mathrm{d}x$$

may not converge. If $f \in \mathcal{L}^1 \cap \mathcal{L}^2$, then it makes sense.

The key idea behind the extension of $\mathcal{F}$ to $\mathcal{L}^2$ is to show that $\mathrm{cl}(S) = \mathcal{L}^2$ (*i.e.* the Schwartz functions are dense in $\mathcal{L}^2(\mathbb{R})$) and the appeal to the following result



Proposition 4.10

Given metric spaces M, N with N complete and $A \subset M$ dense in the metric with $\mathrm{cl}(A) = M$, then given $f : A \to N$ uniformly continuous, there exists a unique continuous $g : M \to N$ with $g\big|_A = f$.

The proof is left as an exercise.

To show that $\mathrm{cl}(\mathcal{S}(\mathbb{R})) = \mathcal{L}^2(\mathbb{R})$ we construct a class of compactly supported smooth functions contained in $\mathcal{S}(\mathbb{R})$. Given $[a, b]$, there exists a smooth function $f \in \mathcal{C}^\infty(\mathbb{R})$ that is of the following form.

This is done via several steps.

Step 1 Construct $f_0 \in \mathcal{C}^\infty(\mathbb{R})$ with

$$f_0(x) = \begin{cases} 0 & \text{if } x \le 0 \\ > 0 & \text{if } x > 0 \end{cases} \quad \text{we use} \quad f_0(x) = \begin{cases} e^{-\frac{1}{x}} & \text{if } x > 0 \\ 0 & \text{if } x \le 0 \end{cases}$$

Claim

$f_0 \in \mathcal{C}^\infty(\mathbb{R})$, but not **real-analytic** at $x = 0$. [11]

We can paste this construction together and define $f_1(x) = f_0(x-a)f_0(b-x)$. Then, clearly,

$$f_1(x) = \begin{cases} 0 & \text{if } x \notin (a, b) \\ > 0 & \text{if } x \in (a, b) \end{cases}$$

To construct $f(x)$ as in the picture, we need an additional step:

[11] Otherwise, the function would have to be zero everywhere.

Lemma 4.11

Given $I = (a,b)$, $\exists f_2 \in \mathcal{C}^\infty(\mathbb{R})$ such that

$$f_2(x) = \begin{cases} 0 & \text{if } x \le a \\ 1 & \text{if } x \ge b \end{cases} \qquad \text{and } 0 < f_2 < 1 \text{ on } (a,b)$$

Proof. Let

$$f_2(x) = \frac{\int_{-\infty}^{x} f_1(\tau)\mathrm{d}\tau}{\int_{-\infty}^{\infty} f_1(\tau)\mathrm{d}\tau}.$$

The proof was not completed in class and is left as an exercise. ■

We have shown that there exists

$$g \in \mathcal{C}^\infty = \begin{cases} 0 & \text{if } x \le a \\ 1 & \text{if } x \ge a + \varepsilon \end{cases}$$

with $0 < g < 1$ on $(a, a+\varepsilon)$ and similarly, there exists $h \in \mathcal{C}^\infty(\mathbb{R})$ with

$$h(x) = \begin{cases} 0 & \text{if } x \le b - \varepsilon \\ 1 & \text{if } x \ge b \end{cases}$$

and $0 < h < 1$ on $(b-\varepsilon, b)$. We define $g(x) = g(x)[1 - h(x)]$ and clearly, $f \in \mathcal{S}(\mathbb{R})$. Recall we want ot show that $\mathrm{cl}(\mathcal{S}(\mathbb{R})) = \mathcal{L}^2(\mathbb{R})$. Here,

$$\mathrm{cl}(\mathcal{S}(\mathbb{R})) = \{f : \|f - f_n\|_{\mathcal{L}^2(\mathbb{R})} \to 0 \text{ for any sequence } \{f_n\} \in \mathcal{S}\}$$

We do this in several steps.

Proposition 4.12

Let $\mathcal{A} = \bigcup_{N=1}^{\infty} I_i$ where I_i are intervals. Then, $\chi_A \in \mathrm{cl}(\mathcal{S}(\mathbb{R}))$.

Proof. Enough to check for $A = [a,b]$ (and extend by linearity) that there exists $f \in \mathcal{S}(\mathbb{R})$ with

$$\int_{\mathbb{R}} |f - \chi_A|^2 \mathrm{d}x < \varepsilon, \qquad \forall\, \varepsilon > 0$$

Choose $f \in \mathcal{C}^\infty(\mathbb{R})$ as in the picture, then $f \equiv \chi_A$ on $(a+\varepsilon, b-\varepsilon)$ and $|f - \chi_A| < 1$ on $[a, a+\varepsilon]$ and $[b-\varepsilon, b]$, therefore $\int_{\mathbb{R}} |f - \chi_A|^2 \mathrm{d}x < 2\varepsilon$ for arbitrary $\varepsilon > 0$. We extend this for measurable sets and simple functions. ■

Proposition 4.13
Let $A \in \mathcal{M}$ with $m(A) < \infty$. Then, $\chi_A \in \mathrm{cl}(\mathcal{S}(\mathbb{R}))$

Proof. Let $\varepsilon > 0$ be given. Then by an ancient result, there exists a finite union of intervals $B = \bigcup_{N=1}^{\infty} I_i$ for I_i intervals such that $m(A\Delta B) < \varepsilon$, where $A\Delta B = (A \setminus B) \cup (B \setminus A)$ and

$$\int_{\mathbb{R}} |\chi_B - \chi_A|^2 \mathrm{d}x = m(A\Delta B) < \varepsilon$$

Then, by linearity, $S_n \in \mathrm{cl}(\mathcal{S}(\mathbb{R}))$ for any simple function S_n. Finally, let $f \geq 0$ and $f \in \mathcal{L}^2(\mathbb{R})$; we showed a long time ago that we can find $\{s_n\} \nearrow f$ as $n \to \infty$ where $s_n \geq 0$ are simple functions. Then,

$$\lim_{n\to\infty} \int_{\mathbb{R}} |s_n - f|^2 \mathrm{d}x = 0$$

by monotone convergence theorem. ■

We have the final theorem

Theorem 4.14
The closure of the Schwartz functions $\mathrm{cl}(\mathcal{S}(\mathbb{R})) = \mathcal{L}^2(\mathbb{R})$; moreover, there exists a unique linear map $\mathcal{F} : \mathcal{L}^2 \to \mathcal{L}^2$ and a unique linear map $\mathcal{F}^{-1}\mathcal{L}^2 \to \mathcal{L}^2$ such that $\mathcal{F}^{-1}\mathcal{F} = \mathrm{I_d}$ and $\|\mathcal{F}f\|_{\mathcal{L}^2}^2 = \|f\|_{\mathcal{L}^2}^2$

Proof. The first part was done before and the second part follows from real analysis extension lemma ■

4.3 Central Limit Theorem

Let $X \subseteq \mathbb{R}^n$ measurable with respect to some measure $m : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ and we assume that $m(X) = 1$ ("m" is a probability measure on Borel subsets $\mathbb{B}$ of $\mathbb{R}^n$.)

Definition 4.15 (Random variable and expectation)

1. A **random variable** $f : X \to \mathbb{R} \cup \{\pm\infty\}$ is a measurable function.
2. The **expectation** (or mean-value) of f is $\mathsf{E}(f) = \int_X f \mathrm{d}m$.

Example 4.4
Let $f = \frac{1}{2}\left(n \sum_{k=1}^{n} R_k\right)$ where the R_k is the k^{th} Rademacher function. Here $X = [0, 1)$ and

$\mathrm{d}m = \mathrm{d}x$. Then

$$\mathsf{E}\left(f\right) = \int_{[0,1]} = \frac{1}{2}\left(n + \sum_{k=1}^{n} R_k\right)\mathrm{d}x = \frac{n}{2}$$

since the Rademacher functions are balanced.

Definition 4.16 (Distributional measure (push-forward measure))

Given a random variable, $f : X \to \mathbb{R} \cup \{\pm\infty\}$ and a Borel subset $A \subseteq \mathbb{R}$, one can define a measure associated to f, m_f on $\mathbb{B}(\mathbb{R})$ as follows:

$$m_f(A) = m(f^{-1}(A))$$

This is called the **push-forward** measure and this specific case the **distributional** measure associated with f.

Exercise 4.1

Check that $m(f^{-1}(A)) = m_f(A)$ is a measure on $\mathbb{B}(\mathbb{R})$.

Let $X \subseteq \mathbb{R}^n$ be Borel measurable with respect to m with $m(X) = 1$ and assume $f : X \to \mathbb{R} \cup \{\pm\infty\}$ random variable. Recall the distribution measure associated with (f, m) is $m_f(A) := m(f^{-1}(A))$ for $A \in \mathbb{B}$; m_f is a measure on $\mathbb{B}$.

Proposition 4.17

Let $\varphi \geq 0$ be Borel measurable function on $\mathbb{R}$. Then

$$\int_{\mathcal{X}} \phi(F)\mathrm{d}m = \int_{\mathbb{R}} \varphi \mathrm{d}m_f$$

Proof. Let $A \in \mathbb{B}$ and consider $\varphi = \chi_A$. The left hand side is $\int_{\mathcal{X}} \chi_A(f)\mathrm{d}m = m(f^{-1}(A))$ by definition. As for the right hand side, $\int_{\mathbb{R}} \chi_A \mathrm{d}m_f = m_f(A)$ and the two are equal by definition of the distribution function. By linearity, the result holds for simple functions s_n.

For general $\varphi \in \mathbb{B}$ for $\varphi \geq 0$, we have simple functions $s_n \varphi$ as $n \to \infty$, therefore $s_n(f) \nearrow \varphi(f)$ as well. If we look at

$$\lim_{n\to\infty} \int s_n(f)\mathrm{d}m = \int_{\mathcal{X}} \varphi(f)\mathrm{d}m$$

by monotone convergence theorem. On the other hand, $s_n(f)\mathrm{d}m = s_n \mathrm{d}m_f$ and

$$\lim_{n\to\infty} \int_R s_n \mathrm{d}m_f \xrightarrow{\text{MCT}} \int_{\mathcal{X}} \varphi \mathrm{d}m_f$$

■

Corollary 4.18

Suppose $\varphi \in \mathbb{B}$. Then $\varphi \in \mathcal{L}1(m_f)$ if and only if $\varphi(f) \in \mathcal{L}^1(m)$.

Definition 4.19

Given random variables $f, g : X \to \mathbb{R} \cup \{\pm\infty\}$ are identically distributed (ID) provided the distribution measures are equal to each other. Now, given random variables $f_1, \ldots, f_n : X \to \mathbb{R} \cup \{\pm\infty\}$. We let $\boldsymbol{f}(x) = (f_1(x), \ldots, f_n(x))$ and $\boldsymbol{f} : X \to \overline{\mathbb{R}^n}$

Definition 4.20

Given $A \in \mathbb{B}(\mathbb{R}^n)$, the **joint probability distribution** $m_{f_1,\ldots,f_n}$ is defined by

$$m_{\boldsymbol{f}}(A) = m(\boldsymbol{f}^{-1}(A))$$

Proposition 4.21

Given $\varphi \in \mathbb{B}(\mathbb{R}^n)$,

$$\int_{\mathcal{X}} \varphi(f_1, \ldots, f_n) \mathrm{d}m = \int_{\mathbb{R}^n} \varphi \mathrm{d}m_{\boldsymbol{f}}$$

The proof is the same as Proposition 4.17

Definition 4.22 (Independence)

The random variables $f_1, \ldots, f_n$ are independent provided if for any $A_1, \ldots, A_n \in \mathbb{B}(\mathbb{R})$,

$$m(f_1^{-1}(A_1) \cap \cdots \cap f_n^{-1}(A_n)) = \prod_{j=1}^{n} m(f_j^{-1}(A_j)) = \prod_{j=1}^{n} m_{f_j}(A_j)$$

Theorem 4.23

$f_1, \ldots, f_n : X \to \mathbb{R}$ are independent if and only if

$$m_{\boldsymbol{f}} = m_{f_1} \times \cdots \times m_{f_n}$$

the product measure.

Proof. (Sketch) Let's check this for product sets of the form $A = A_1 \times \cdots \times A_n$.

$$m_{\boldsymbol{f}}(A_1 \times \cdots \times A_n) = m(\boldsymbol{f}^{-1}(A)) = m(f_1^{-1}(A_1) \cap \cdots \cap f_n^{-1}(A_n)) = \prod_{i=1}^{n} m_{f_i}(A_i)$$

■

Remark

Suppose $f_1, \dots, f_n$ are independent and integrable. Then

$$\int_{\mathcal{X}} f_1 \times \cdots, \times f_n \mathrm{d}m = \left(\int_{\mathcal{X}} f_1 \mathrm{d}m\right) \times \cdots \times \left(\int_{\mathcal{X}} f_n \mathrm{d}m\right)$$

Proof. By Proposition 4.21, where $f_1 = \varphi(f_1)$ and $\varphi(x_1) = x_1$

$$\begin{aligned}
\int_{\mathcal{X}} f_1 \times \cdots \times f_n \mathrm{d}m &= \int_{\mathbb{R}^n} x_1 \cdots x_n \mathrm{d}m_{\boldsymbol{f}} \\
&= \int_{\mathbb{R}^n} x_1 \cdots x_n \mathrm{d}m_{f_1} \times \cdots \times \mathrm{d}m_{f_n} && \text{(by Theorem 4.23)} \\
&= \prod_{i=1}^{n} \left(\int_{\mathbb{R}} x_i \mathrm{d}m_{f_i}\right) && \text{(by Fubini)}
\end{aligned}$$

■

Theorem 4.24 (Law of large numbers)

Let $f_1, \dots, f_n : X \to \mathbb{R}$ be bounded independently and identically distributed random variables (IID) with $\mathsf{E} = \mathsf{E}(f_i) = \int_{\mathcal{X}} f_i \mathrm{d}m$ for $i = 1, \dots, n$. Then

$$m\left(\left\{x \in X; \lim_{n\to\infty} \frac{f_1(x) + \cdots + f_n(x)}{n} \xrightarrow{n\to\infty} \mathsf{E}\right\}\right)$$

Proof. Essentially identical to Rademacher case (exercise), *i.e.* we consider $S_n(x) = f_1(x) + \cdots + f_n(x) - n\mathsf{E}$ and $\frac{S_n(x)}{n} \to 0$ as $n \to \infty$. Of one checks the proof carefully (look at Rademacher),

$$m\left(\left\{x \in X; \frac{S_n(x)}{n^{\frac{1}{2}+\alpha}} \text{ for fixed } \alpha > 0\right\}\right) = 1$$

■

The Central Limit Theorem is concerned with the asymptotic behavior as $n \to \infty$ of $\frac{S_n(x)}{\sqrt{n}}$. Given $f_1, \dots, f_n : X \to \mathbb{R}$ IID random variables, we define the **variance** by

$$\mathsf{Var}(f) = \int_{\mathcal{X}} (f_i - \mathsf{E}(f_i))^2 \mathrm{d}m$$

Theorem 4.25 (Central Limit Theorem)

Suppose $f_1, \dots, f_n$ are IID random variables on a probability space $\mathcal{X}$ *i.e.* $m(x) = 1$ and such that $|f_j(x)| \leq M \; \forall\, x \in \mathcal{X}$, that is they are bounded. Write $S_n(x) = f_1(x) + \cdots + f_n(x) - n\mathsf{E}(f_i)$ where $\mathsf{E}(f_i) = \int_{\mathcal{X}} f_i \mathrm{d}m$. Wlog, assume $\mathsf{E}(f_i) = 0$, by shifting $f_j \mapsto f_j - \mathsf{E}(f_j)$, and f_i^2 bounded with $\sigma^2 = \mathsf{Var}(f_i) < \infty$. Let m_n be the distribution measure of $s_n/\sqrt{n}$ and define $m_n \coloneqq m_{S_n}/\sqrt{n}$ for $J = (a, b) \subset \mathbb{R}$. Then, for any $a, b \in \mathbb{R}$ with $a < b$, then

$$\lim_{n\to\infty} m\left(\left\{x \in X \,\middle|\, a < \frac{S_n}{\sqrt{n}} < b\right\}\right) = \frac{1}{\sqrt{2\pi}\sigma}\int_a^b \exp\left(-\frac{t^2}{2\sigma^2}\right)\mathrm{d}t \tag{4.60}$$

where $e^{-\frac{t^2}{2}}$ is the Gaussian distribution. The LHS in (4.60) is $m_n(J)$, while the right hand side is $m_{\sigma^2}(J) = \frac{1}{\sqrt{2\pi}\sigma}\int_J e^{-\frac{t^2}{2}}\mathrm{d}t$

Note

The Central limit theorem simply says that $\lim_{n\to\infty} m_n(J) = m_{\sigma^2}(J)$ for any interval $J = (a, b)$, that is weak limit is m_{σ^2}

Proof. We take Fourier transform of m_n and m_{σ^2}

Lemma 4.26

Let $\chi_n(t) = \int_{\mathbb{R}} e^{-ixt}\mathrm{d}m_n(x)$. Then, for any $t \in \mathbb{R}$,

$$\lim_{n\to\infty} \chi_n(t) = e^{-\frac{\sigma^2 t^2}{2}} \tag{4.61}$$

Proof.

$$\begin{aligned}
\chi_n(t) &= \int_{\mathbb{R}} e^{-xt}\mathrm{d}m_n(t) \\
&= \int_{\mathcal{X}} e^{-it\left(\frac{S_n}{\sqrt{n}}\right)}\mathrm{d}m \\
&= \int_{\mathcal{X}} e^{-it\left(\frac{f_1+\cdots+f_n}{\sqrt{n}}\right)}\mathrm{d}m && (\text{ as } t = 0) \\
&= \int_{\mathcal{X}} \prod_{j=1}^{n}\left(e^{-it\frac{f_j}{\sqrt{n}}}\right)\mathrm{d}m \\
&= \prod_{j=1}^{n}\int_{\mathcal{X}} e^{-it\frac{f_j}{\sqrt{n}}}\mathrm{d}m && (\text{by independence}) \\
&= \left(\int_{\mathcal{X}} e^{-it\frac{f_j}{\sqrt{n}}}\mathrm{d}m\right)^n && (\; f_j \text{ are identically distributed})
\end{aligned}$$

assuming that f is bounded to simplify life. Now by Taylor expansion,

$$e^{-it\frac{f}{\sqrt{n}}} = 1 - \frac{itf}{\sqrt{n}} - \frac{t^2}{\sqrt{n}} - \frac{t^2}{2n}f^2(1 + r_n)$$

where $|r_n(t)| \leq M_n$ for all $t \in \mathbb{R}$ and $r_n \to 0$ as $n \to \infty$. We then have

$$\begin{aligned} \Rightarrow \left(\int_{\mathcal{X}} e^{-it\frac{f_j}{\sqrt{n}}} \mathrm{d}m \right)^n &= \left(\int_{\mathcal{X}} \left[1 - \frac{itf}{\sqrt{n}} - \frac{t^2}{\sqrt{n}} - \frac{t^2}{2n} f^2 (1 + r_n) \right] \mathrm{d}m \right)^n \\ &= \left(1 - \left(\frac{t^2 \sigma^2}{2n} \right) (1 + \varepsilon_n) \right)^n \end{aligned}$$

since $\sigma^2 = \int_{\mathcal{X}} f^2$ where $\varepsilon_n \to 0^+$ and as $n \to \infty$, this converges to $e^{-\frac{t^2\sigma^2}{2}}$ (to see this, take logarithms). ■

Claim

Given any $f \in \mathcal{S}(\mathbb{R})$,

$$\lim_{n\to\infty} \int_R f \mathrm{d}m_n = \int_{\mathbb{R}} f \mathrm{d}m_{\sigma^2}$$

Proof. We use Fourier inversion and Plancherel:

$$\begin{aligned} \int_{\mathbb{R}} f \mathrm{d}m_n &= \frac{1}{2\pi} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \hat{f}(t) e^{ixt} \mathrm{d}t \right) \mathrm{d}m_n(x) && \text{(by Fourier inversion)} \\ &= \lim_{n\to\infty} \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(t) \chi_n(t) \mathrm{d}t && \text{(by Fubini)} \\ &= \int_{\mathbb{R}} \hat{f}(t) e^{-\frac{t^2\sigma^2}{2}} \mathrm{d}t \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} f(t) e^{-\frac{t^2}{2\sigma^2}} \mathrm{d}t \end{aligned}$$

■

The last step consists in approximating $\chi_J = \chi_{(a,b)}$ by $f_\varepsilon \in \mathcal{S}(\mathbb{R})$ and apply dominated convergence theorem to show that $\lim_{n\to\infty} \chi_n(J) = \chi_{\sigma^2}(J)$ (exercise). ■

Example 4.5

Let $X = [0, 1]$ and m the Lebesgue measure, with $f_n = R_j$. By the strong law of large number, $S_n(\omega)/n \xrightarrow{\text{a.s.}} 0$ for $\omega \in [0, 1]$. The question one might ask is how many trials does it take to be reasonably sure that S_n/n is near zero ? For example, we want to know that $S_n/n < 0.01$ with probability 99%. First, its easy to check that the variance of the

Rademacher function is $\sigma^2 = \mathsf{Var}\,(R_i) = 1$, then

$$\begin{aligned} 0.99 &= m\left(\left\{\omega \in [0,1] : \frac{|S_n|}{n} < 0.01\right\}\right) \\ &= m\left(\left\{\omega \in [0,1] : \frac{|S_n|}{\sqrt{n}} < 0.01\sqrt{n}\right\}\right) \\ &\sim_{n\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-0.01\sqrt{n}}^{0.01\sqrt{n}} e^{-\frac{t^2}{2}}\,\mathrm{d}t \end{aligned}$$

and numerically, $\sqrt{n}0.01 \approx 2.57$ and so we need $n \approx 66,000$.

Index