Workbook of Real Analysis

Based on the Textbook: Princeton Lectures in Anaylsis III

Author: Ding Rui

Date: September 1, 2021

Version: 1.0

ElegantLATEX Program

Contents

Chapter 1 Measure Theory

Key words

- ❏ *Exterior measure*
- ❏ *(Lebesgue) measurable set*
- ❏ *(Lebesgue) measurable functions*
- ❏ *σ-algebra and Borel sets*
- ❏ *Littlewood's three principles*
- ❏ *The Cantor set*

1.1 Exercises

Exercise 1.9 (Thick Boundary) Give an example of an open set $\mathcal{O}$ with $m(\partial\mathcal{O}) > 0$.

Hint Enumerate rationals in $[0, 1]$ and denote the n-th rationals as r_n. Then $B = \bigcup B(r_n, 1/2^{n+1})$ is what we desire.

Note *It is intriguing that B^c is a closed set with $m(B^c \cap [0, 1]) > 0$, but $B^c \cap [0, 1]$ contains no rationals.*

Exercise 1.13 The following deals with G_δ and F_σ sets.

(a) Show that a closed set is a G_δ and an open set an F_σ.

(b) Give an example of an F_σ which is not an G_δ.

(c) Give an example of a Borel set which is not a G_δ nor an F_σ.

Solution

(a) If F is closed, we consider the open sets $\mathcal{O}_n = \{x : d(x, F) < 1/n\}$. In fact, $F = \bigcap \mathcal{O}_n$. As for the open set, take the complement.

(b) Let F be a denumerable set that is dense at the same time. Clearly F is an F_σ [Say, $\mathbb{Q}$]. Assume $F = \bigcap \mathcal{O}_n$. If we subtract the n-th element f_n of F from $\mathcal{O}_n$, then we must have $\emptyset = \bigcap(\mathcal{O}_n - \{f_n\})$. To avoid notational clutter, let F_n denote $\bigcap_{k\le n}(\mathcal{O}_k - \{f_k\})$, then F_n is open and dense for all n. Moreover, $F_n \searrow \emptyset$.

Now we may choose $[a_0, b_0] \subseteq F_0$. Since F_1 is dense, we can always select a sub-interval $[a_1, b_1] \subseteq F_1 \cap [a_0, b_0]$. Proceed this procedure indefinitely, and we will get a sequence $[a_{i+1}, b_{i+1}] \subseteq F_{i+1} \cap [a_i, b_i]$. Note at this point, however, that the nested intervals theorem guarantees $\bigcap[a_i, b_i] \neq \emptyset$, which contradicts the property of $\{F_n\}$ aforementioned.

(c) For $x < 0$ let all rationals in F. Otherwise, let all irrationals in F.

■

Exercise 1.23 (Separate Continuity) Suppose $f(x, y)$ is a function on $\mathbb{R}^2$ that is separately continuous: for each fixed variable, f is continuous in the other variable. Prove that f is measurable.

Solution Define $f_n(\frac{k+\eta}{2^n}, y) = (1-\eta)f(\frac{k}{2^n}, y) + \eta f(\frac{k+1}{2^n}, y)$, for $k \in \mathbb{N}$ and $\eta \in [0, 1)$. Clearly, f_n is measurable and $f_n \to f$ for all (x, y). ■

Remark Moreover, for any $\varepsilon > 0$, there exists a closed set F such that $m(F^c) < \varepsilon$ and $f\big|_F$ is continuous.

Exercise 1.35' Show that not all Lebesgue measurable sets are Borel sets from these two perspectives:
(a) Continuous functions.
(b) Slicing.

Hint For (a), you may want to show that some continuous functions map a measurable set to a non-measurable one, while all the continuous functions map Borel sets to Borel sets. For (b), you may want to show that there exists a measurable set such that some of its slices can be non-measurable, while the slices of Borel sets are still Borel. To deal with the Borel set, you may find **good set principle** rather helpful.

1.2 Problems

Problem 1.4 Prove that a bounded function on an interval $J = [a, b]$ is Riemann integrable iff its set of discontinuities has measure 0.

Proceed in this way:
(a) For every $\varepsilon > 0$, the set of points c in J s.t. $\text{osc}(f, c) \geq \varepsilon$ is compact.
(b) Establish the sufficiency.
(c) Establish the necessity.

Solution
(a) It is not hard to observe.
(b) Given $\varepsilon > 0$, let $A_\varepsilon = \{c \in J : \text{osc}(f, c) \geq \varepsilon\}$, then $m(A_\varepsilon) = 0$. Let $\mathcal{A}$ be an open coverage whose measure is less than ε. By virtue of Heine-Borel property, we can pick an coverage comprised of only a finite number of intervals. Denote its number to be N. We may assume that $\mathcal{A}$ is union of finitely many intervals.
On compact set $J - \bigcup_{I \in \mathcal{A}} I$, function f is continuous. Thus, it is uniformly continuous. I.e., there exists δ such that $|x - y| < \delta$ implies $|f(x) - f(y)| < \varepsilon$. As long as the partition is less than $\min(\varepsilon, \delta)$, at most $3N$ intervals will intersect with A_ε. The difference of this partition is at most $m([a, b])\varepsilon + 6NM\varepsilon$, where $M < \infty$ satisfies that $M > |f|$. Since N is predetermined by the coverage, the bound can be arbitrarily close to 0.
(c) We show that $A_{1/n}$ has measure zero. For any given $\varepsilon > 0$, there exists a partition such that its difference is less than $\frac{\varepsilon}{n}$. If one of its interior contains a point in $A_{1/n}$, the amptitude of this interval is at least $\frac{1}{n}$. By Chebyshev inequality, we have that the length of such intervals is less than ε. Futhermore, the boundary has measure 0, so $A_{1/n}$ can be covered by an open set of measure ε. We complete the proof.

■

Remark The proof above relies heavily on the simple structure of open sets and partition in $\mathbb{R}$. To generalize it to $\mathbb{R}^d$, you may find that Exercise 1.13 (a) comes in handy. Also, you may try to substitute J with a compact set.

Problem 1.5 Suppose E is measurable with finite measure, and

$$E = E_1 \cup E_2, \quad E_1 \cap E_2 = \emptyset.$$

If $m(E) = m_*(E_1) + m_*(E_2)$, then both E_1 and E_2 are measurable.

Solution We may assume that E is open. [Exercise: Generalize the following to any E.]

It suffices to show that for any given $\varepsilon > 0$, there exists an open set $E_1' \supseteq E_1$ such that $m_*(E_1' - E_1) < \varepsilon$. Indeed, there exists an open set E_1' that is sandwiched between E_1 and E. Moreover, $m(E_1') - m_*(E_1) < \frac{\varepsilon}{2}$. Likewise, we can find E_2'. We have

$$\begin{aligned} m_*(E_1' - E_1) &\le m\left((E_1' - E_1) \cup (E_2' - E_2)\right) \\ &= m(E - (E - E_1') \cup (E - E_2')) \\ &= m(E) - m(E - E_1') - m(E - E_2') \\ &= m(E_1') + m(E_2') - m(E) < \varepsilon, \end{aligned}$$

which gives the desired result. ■

Problem 1.5' Suppose E is measurable with finite measure, then for any subset $F \subseteq \mathbb{R}^d$ that is disjoint with E, the identity $m_*(E \cup F) = m(E) + m_*(F)$ always holds.

Solution WLOG, assume that $m_*(G) < \infty$, where G denotes $E \cup F$. Given any $\varepsilon > 0$, we may choose a closed set $E' \subseteq E$ with $m(E - E') < \frac{\varepsilon}{2}$ and an open set $G' \supseteq G$ with $m(G') - m_*(G) < \frac{\varepsilon}{2}$. Note that $G' - E'$ is open. Besides, $F \subseteq G' - E'$. Finally, $m_*(F') \le m(G' - E') = m(G') - m(E') < m_*(G) - m(E) + \varepsilon$. Since ε can be arbitrarily small, $m_*(G) \ge m(E) + m_*(F)$. ■

1.3 Supplementaries

Exercise 1.a In this exercise, we show that several other operations perserve measurability while some may not. We begin by assume f to be measurable on $\mathbb{R}^d$.

(a) Show that

$$g(x) = \limsup_{r \searrow 0} |f(x+r) - f(x)|$$

is measurable.

(b) Further assume that f is continuous. Prove that

$$h(x) = \limsup_{r \searrow 0} \left| \frac{f(x+r) - f(x)}{r} \right|$$

is also measurable.

Solution For (a), note that $\overline{f}_r(x) = \sup_{y \in [0,r]} f(x+y)$ is measurable due to the fact that $\bigcup_{0 \le y \le r} (E+y)$ is measurable as long as E is measurable. So is $\underline{f}_r(x) = \inf_{y \in [0,r]} f(x+y)$. Then $f_r(x) = \sup_{y \in [0,r]} |f(x+y) - f(x)|$ is clearly measurable since

$$f_r(x) = \max(\overline{f}_r(x) - f(x), f(x) - \underline{f}_r(x)).$$

By a limiting argument, clearly (a) holds.

For (b), we approach $g_r(x) = \sup_{y\in[0,r]} \left|\frac{f(x+y)-f(x)}{y}\right|$ by taking supreme over all functions defined by $\left|\frac{f(x+y)-f(x)}{y}\right|$ where $y \in \mathbb{Q}$ and is between 0 and r. It follows that $h = \lim_{\substack{n\to\infty \\ n\in\mathbb{N}}} g_{1/n}$ is measurable. ■

Exercise 1.b (Carathéodory Measurability) A set $E \subseteq \mathbb{R}^d$ is **Carathéodory measurable** if for every $A \subseteq \mathbb{R}^d$,

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

This condition is sometimes referred to as the separation condition. Prove that this notion coincides with Lebesgue measurability in $\mathbb{R}^d$.

In fact, this ingenious observation first made by Carathéodory generalizes the notion of measurability given an outer measure. Refer to Chapter 7 for more information.

Solution Let E be Lebesgue measurable. Find a sequence of open sets $\mathcal{O}_n \searrow \mathcal{O}$ with the property that $A \subseteq \mathcal{O}$ and that $m_*(A) = m(\mathcal{O})$. We have for every $n \in \mathbb{N}$, $m_*(\mathcal{O}_n) = m_*(E \cap \mathcal{O}_n) + m_*(E^c \cap \mathcal{O}_n)$. Letting n tends to ∞, we have

$$m_*(A) = m(\mathcal{O}) = m_*(E \cap \mathcal{O}) + m_*(E^c \cap \mathcal{O}) \geq m_*(E \cap A) + m_*(E^c \cap A).$$

The other side of inequality is a simple argument of sub-additivity. Combined these two sides, E is Carathéodory measurable.

Now we turn to the reverse order. Taking $A = C_n$ as the cube centered at the origin and with side length n, we have $E \cap C_n$ is Lebesgue measurable due to Problem 1.5 above. Consequently, $E = \bigcup_n (E \cap C_n)$ is Lebesgue measurable. ■

Chapter 2 Integration Theory

Key words

- *Simple functions*
- *Step functions*
- *Lebesgue integrability*
- *Bounded convergence theorem*
- *Monotone convergence theorem*
- *Dominated convergence theorem*
- *Completeness of the space $L_1(\mathbb{R}^d)$ of integrable functions*
- *Invariance properties*
- *Fubini's theorem*
- *Convolution*
- *Approximations to the identity*

2.1 Exercises

Exercise 2.7 Let $\Gamma \subseteq \mathbb{R}^d \times \mathbb{R}, \Gamma = \{(x, y) \in \mathbb{R}^d \times \mathbb{R} : y = f(x)\}$, and assume f is measurable. Show that Γ is a measurable subset of $\mathbb{R}^{d+1}$, and $m(\Gamma) = 0$.

Solution Let $F(x, y) = f(x) - y$, which clearly is measurable. In particular, $F^{-1}(0) = \Gamma$ is measurable. Hence $m(\Gamma) = \int_{\mathbb{R}^d} \left(\int \Gamma_x \,\mathrm{d}y\right) \,\mathrm{d}x = 0$. ■

Exercise 2.8 Suppose f is integrable on $\mathbb{R}$, then the indefinite integral $F(x) = \int_{\infty}^{x} f(y) \,\mathrm{d}y$ is absolutely continuous.

Just to emphasize this property.

Exercise 2.15 Consider the function defined over $\mathbb{R}$ by

$$f(x) = \begin{cases} x^{-1/2} & \text{if } 0 < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

For a fixed enumeration $\{r_n\}_{n=1}^{\infty}$ of $\mathbb{Q}$, let

$$F(x) = \sum_{n=1}^{\infty} 2^{-n} f(x - r_n).$$

Show some properties of F. [omitted]

Solution Clearly F is measurable and non-negative. We have

$$\int F = \sum_{n=1}^{\infty} \int 2^{-n} f(x - r_n) \,\mathrm{d}x = 2.$$

Thus, F is integrable and F converges a.e.. Fixed any rational r_n, we see that F is unbounded in the vicinity of r_n. Due to the density of rantional numbers, F is surely unbounded on every interval. ■

Exercise 2.19 Suppose f is integrable on $\mathbb{R}^d$. For each $\alpha > 0$, Let $E_\alpha = \{x : f(x) > \alpha\}$. Prove that

$$\int_{\mathbb{R}^d} |f(x)| \, \mathrm{d}x = \int_0^\infty m(E_\alpha) \, \mathrm{d}\alpha.$$

Solution Let $A \subseteq \mathbb{R}^{d+1}$ such that $A = \{(x, y) : 0 \leq y \leq |f(x)|\}$. Then A is measurable and we have

$$\int_{\mathbb{R}^d} |f(x)| \, \mathrm{d}x = m(A) = \int_0^\infty m(E_\alpha) \, \mathrm{d}\alpha,$$

where the last identity is ensured by Fubini's theorem. ■

Exercise 2.24 Consider the convolution

$$(f * g)(x) = \int_{\mathbb{R}^d} f(x - y) g(y) \, \mathrm{d}y.$$

(a) Show that $f * g$ is uniformly continuous when f is integrable and g is bounded.

(b) If in addition g is integrable, prove that $(f * g)(x) \to 0$ as $|x| \to \infty$.

(c) Show that the bounded property of g is neccesary for (b) to hold.

Solution

(a) Suppose $M = \sup |g| < \infty$. Given $\varepsilon > 0$, since

$$\int_{\mathbb{R}^d} |f(x + h) - f(x)| \, \mathrm{d}x \to 0, \text{ as } h \to 0,$$

we have

$$\begin{aligned}(f * g)(x + h) - (f * g)(x) &= \int_{\mathbb{R}^d} |f(x + h - y) - f(x - y)| \, |g(y)| \, \mathrm{d}y \\ &\leq M \int_{\mathbb{R}^d} |f(y + h) - f(y)| \, \mathrm{d}y,\end{aligned}$$

whose convergence is bounded in a way independent of x. Hence $f * g$ is uniformly continuous.

(b) Clearly $f * g$ is integrable. Combined with the fact that $f * g$ is uniformly continuous,

$$\lim_{|x| \to \infty} (f * g)(x) = 0$$

must hold.

(c) You may find that the functions F and f defined in Exercise 2.15 work here.

■

2.2 Problems

Problem 2.3' A sequence $\{f_n\}$ of measurable functions on $\mathbb{R}^d$ is **Cauchy in measure** if for every $\varepsilon > 0$,

$$m(\{x : |f_k(x) - f_l(x)| > \varepsilon\}) \to 0 \quad \text{as } k, l \to \infty.$$

We say that $\{f_n\}$ **converges in measure** to a (measurable) function f if for every $\varepsilon > 0$

$$m(\{x : |f_k(x) - f(x)| > \varepsilon\}) \to 0 \quad \text{as } k \to \infty.$$

Note that this notion coincides with the "convergence in probability" of probability theory. In the following, we delve into the relationship between different notions of convergences we have encountered so far.

(a) Show that if $\{f_n\}$ is bounded by an integrable function g on $\mathbb{R}^d$ and is Cauchy in measure, then $\{f_n\}$ converges in L_1.
(b) A Cauchy sequence in measure converges in measure.
(c) In space $L_1(\mathbb{R}^d)$, convergence in L_1 implies convergence in measure.
(d) Give an example to show that the notions of convergence in measure and in L_1 do not agree.

Solution The core observation is as follows.

Lemma 2.1 (Completeness in measure)

If a sequence $\{f_n\}$ of measurable functions is Cauchy in measure, there exists a sub-sequence $\{f_{n_k}\}$ such that it converges pointwise to f almost everywhere. Moreover, $f_{n_k} \to f$ in measure. ♡

Proof Pick an increasing sequence $\{n_k\}_{k=1}^{\infty}$ which satisfies

$$m(\{x : |f_{n_l}(x) - f_{n_r}(x)| > 2^{-k-1}\}) < 2^{-k-1} \quad \text{for all } l, r \geq k.$$

Let

$$A_k = \left\{ x : \left| \sup_{l \geq k} f_{n_l}(x) - \inf_{l \geq k} f_{n_l}(x) \right| > 2^{-k} \right\} \quad \text{and} \quad A = \left\{ x : \limsup_{l} f_{n_l}(x) \neq \liminf_{l} f_{n_l}(x) \right\}.$$

Clearly $m(A_k) \leq 2^{-k}$. Moreover, $A \subseteq \liminf A_k$. By the Borel-Cantelli lemma, $m(A) = 0$. This fact directly implies that $f_n \to f$ pointwise almost everywhere.

To prove the second half of the lemma, we introduce the following notations:

$$B_k^{\varepsilon} = \{x : |f_k(x) - f(x)| > 2\varepsilon\}.$$

Several observations come in order. First, for sufficiently large k, $B_{n_k}^{\varepsilon} \subseteq A_k$, which indicates that $f_{n_k} \to f$ in measure. Second, one might observe that $B_k \subseteq \{x : |f_k(x) - f_{n_r}(x)| > \varepsilon\} \cup B_{n_r}^{\varepsilon}$, where $n_r = \min\{n_l : n_l \geq k\}$. As $k \to \infty$, both of the sets on the right hand side tend to 0 in measure. The lemma is proved. □

Back to the original problem.

(a) For any $\varepsilon > 0$, select a ball $B = \{|x| \leq C\}$ so large that $\int_{B^c} g < \varepsilon$. Also, there is a $\delta > 0$ such that $\int_E g < \varepsilon$ whenever $m(E) < \delta$. Select an n sufficiently large such that

$$m\left(\left\{ x : |f_n(x) - f(x)| > \frac{\varepsilon}{m(B)} \right\} \right) < \delta.$$

For now let A denote the set in the measure function. We have

$$\begin{aligned} \int_{\mathbb{R}^d} |f_n(x) - f(x)| \, \mathrm{d}x &= \int_{B^c} |f_n(x) - f(x)| \, \mathrm{d}x + \int_{B \cap A} |f_n(x) - f(x)| \, \mathrm{d}x + \int_{B - A} |f_n(x) - f(x)| \, \mathrm{d}x \\ &\leq 2\varepsilon + \varepsilon + \frac{\varepsilon}{m(B)} \cdot m(B) = 4\varepsilon. \end{aligned}$$

Hence $f_n \to f$ in L_1.

(b) As is shown in the lemma above.
(c) Use Chebyshev inequality.
(d) Let

$$f_n = \begin{cases} n & \text{if } 0 < x < 1/n, \\ 0 & \text{otherwise,} \end{cases}$$

and $f \equiv 0$. The rest should not be hard to work out.

■

Problem 2.5 Under the premise of admitting **continuum hypothesis**, construct an ordering $\prec$ of $\mathbb{R}$ such that for each $y \in \mathbb{R}$ the set $\{x \in \mathbb{R} : x \prec y\}$ is at most countable.

Solution First select a well-ordering $\prec$ of $\mathbb{R}$, whose existence is guaranteed by the long established well-ordering principle. Define the set $X = \{y \in \mathbb{R} : \text{the set } \{x \in \mathbb{R} : x \prec y\} \text{ is not countable}\}$. If $X = \emptyset$ we are done. Otherwise, let $\tilde{y}$ be the minimum of X. Since the cardinality of the set $\{x \in \mathbb{R} : x \prec \tilde{y}\}$ is exactly $\aleph_1$, we can build a bijection from this set to $\mathbb{R}$. Hence a new ordering $\tilde{\prec}$ of $\mathbb{R}$ can be constructed. Note that the well-ordering property is preserved under this bijection and that $\tilde{\prec}$ satisfies all the conditions. ■

Chapter 3 Differentiation and Integration

Key words

- *Maximal function*
- *Covering lemmas*
- *The Lebesgue differentiation theorem*
- *Points of Lebesgue density*
- *Lebesgue set (of a function)*
- *Bounded eccentricity*
- *Approximations to the identity (w.r.t. convolution kernels)*
- *Rectifiability (of curves)*
- *Bounded variation*
- *Analytic properties of functions of bounded variation*
- *Absolute continuity*
- *Newton-Leibniz formula*
- *Minkowski content*
- *Isoperimetric inequality*

3.1 Exercises

Exercise 3.1 Suppose φ is an integrable function on $\mathbb{R}^d$ with $\int_{\mathbb{R}^d} \varphi = 1$. Set $K_\delta(x) = \delta^{-d}\varphi(x/\delta)$ for $\delta > 0$.

(a) $\{K_\delta\}_{\delta>0}$ is a family of good kernels.

(b) Assume in addition that φ is bounded and supported in a bounded set. Verify that $\{K_\delta\}_{\delta>0}$ is an approximation to the identity.

(c) Suppose that f is integrable on $\mathbb{R}^d$. Clearly $f * K_\delta$ is integrable. Moreover,

$$\|(f * K_\delta) - f\|_{L^1(\mathbb{R}^d)} \to 0 \quad \text{as } \delta \to 0.$$

Here we only assume K_δ to be good kernels, regardless of the particular form in the description part.

Solution (a) and (b) shall not be baffling. Here we only focus on (c).

For any fixed $\varepsilon > 0$, select $\eta > 0$ so small that whenever $|h| < \eta$ there is $\int_{\mathbb{R}^d} |f(x+h) - f(x)| \,\mathrm{d}x < \varepsilon$. Meanwhile, select δ so small that $\int_{|x|\geq\eta} K_\delta < \varepsilon$. Denote M as the supremum of $\int_{\mathbb{R}^d} |f|$ and $\int_{\mathbb{R}^d} |K_\delta|$ for $\delta > 0$. It holds that

$$\begin{aligned}
\|(f * K_\delta) - f\|_{L^1(\mathbb{R}^d)} &\leq \int_{\mathbb{R}^d} \left(\int_{|y|\geq\eta} + \int_{|y|<\eta} \right) |f(x-y) - f(x)| \cdot |K_\delta(y)| \,\mathrm{d}y \,\mathrm{d}x \\
&= \left(\int_{|y|\geq\eta} + \int_{|y|<\eta} \right) |K_\delta(y)| \cdot \int_{\mathbb{R}^d} |f(x-y) - f(x)| \,\mathrm{d}x \,\mathrm{d}y \\
&\leq 2M\varepsilon + M\varepsilon = 3M\varepsilon.
\end{aligned}$$

■

Exercise 3.2' The following deals with the maximal function f^*.

(a) Show that f^* may discontinue at some points in $\mathbb{R}^d$. Two possibilities can be identified at the discontinuities x: either $f^*(x) = \infty$ or $f^*(x) \leq \liminf_{|r|\searrow 0} f^*(x+r)$.

(b) If f is integrable on $\mathbb{R}^d$ and is not identically zero, then f^* is not integrable on $\mathbb{R}^d$.

Exercise 3.6 (One-sided maximal function) In one dimension we define the "one-sided" maximal function

$$f_+^*(x) = \sup_{h>0} \frac{1}{h} \int_x^{x+h} |f(y)| \, \mathrm{d}y.$$

If $E_\alpha^+ = \{x \in \mathbb{R} : f_+^*(x) > \alpha\}$, then

$$m(E_\alpha^+) = \frac{1}{\alpha} \int_{E_\alpha^+} |f(y)| \, \mathrm{d}y \tag{3.1}$$

for all integrable functions f on $\mathbb{R}$.

Solution It is more intuitive to view the identity as

$$\alpha = \frac{1}{m(E_\alpha^+)} \int_{E_\alpha^+} |f(y)| \, \mathrm{d}y.$$

Define $F(x) = \int_{-\infty}^x |f(y)| \, \mathrm{d}y - \alpha x$. Then $x \in E_\alpha^+$ is equivalent to the property that $F(x) < F(y)$ for some $y > x$. Note that E_α^+ is open, and each interval in its decomposition is bounded. By the rising sun lemma, the conclusion holds on each of these intervals, and surely the equation 3.1 holds for $m(E_\alpha^+)$. ■

Exercise 3.9 Let F be a measurable set in $\mathbb{R}^d$ and $\delta(x)$ the distance from x to F, i.e., $\delta(x) = d(x, F)$. Clearly, $\delta(x + y) \leq |y|$ whenever $x \in F$. Prove the more refined estimate: $\delta(x + y) = o(|y|)$ for a.e. $x \in F$.

Exercise 3.7 If a measurable subset E of $[0, 1]$ satisfies $m(E \cap I) \geq \alpha \cdot m(I)$ for some $\alpha > 0$ and all intervals I in $[0, 1]$, then $m(E) = 1$.

Hint The case where $\alpha = 1$ is trivial. Otherwise, consider $[0, 1] - E$. What happens if $m([0, 1] - E) > 0$?

Exercise 3.11 If $a, b > 0$, let

$$f(x) = \begin{cases} x^a \sin(x^{-b}) & \text{for } 0 < x < 1, \\ 0 & \text{if } x = 0. \end{cases}$$

Prove that f is of bounded variation in $[0, 1]$ iff $a > b$. Then, by taking $a = b$, construct (for each $0 < \alpha < 1$) a function that satisfies the Lipschitz condition of exponent α:

$$|f(x) - f(y)| \leq A|x - y|^\alpha$$

but which is not of bounded variation.

Solution As the first step, we show that if $a > b$ then f is of bounded variation on $[c, 1]$ whenever $c > 0$. Taken derivative we have

$$f'(x) = ax^{a-1} \sin(x^{-b}) - bx^{a-b-1} \cos(x^{-b}) \quad \text{for } x \in (0, 1],$$

whose absolute is (Riemann) integrable. By invoking Darboux's theorem, we have that variation goes to 0 as the diameter of the partition on $[c, 1]$ grows to 0. For any partition of $[c, 1]$ fine enough, we have

$$\sum_{i=1}^N |f(x_i) - f(x_{i-1})| = \sum_{i=1}^N |f'(\xi_i)| \Delta x_i \leq \int_0^1 |f'(x)| \, \mathrm{d}x + 1.$$

Since refinement only amplifies the total variation, the claim has been verified. Moreover, the bound on the right hand side is universal for all $c \in [0, 1)$. This fact immediately leads to the bounded variation property of f. The explanation is as follows: for any given partition on $[0, 1]$ with $x_0 = 0$, its total variation is at most

$$|f(x_1) - f(0)| + 1 + \int_0^1 |f'(x)| \, \mathrm{d}x \le 2 + \int_0^1 |f'(x)| \, \mathrm{d}x.$$

Now we turn to the case where $b \ge a > 0$. We construct a partition with $x_0 = 1$ and $x_n = (\pi/2 + n\pi)^{-1/b}$. Given an upper bound N the total variation of partition formed by $x_0, \ldots, x_N, 0$ is at least

$$\sum_{i=2}^{N} \left(\frac{\pi}{2} + i\pi\right)^{-a/b},$$

which grows to infinity as $N \to \infty$. Hence f cannot be of bounded variation. ■

To construct such functions, let $a = b = \alpha$. It can be easily verified that

$$g(x) = \begin{cases} \alpha \frac{\cos(x^{-\alpha})}{x} & \text{if } x \in (0, 1), \\ 0 & \text{otherwise.} \end{cases}$$

is integrable. Moreover, it can be verified that $G(x) = \int_{-\infty}^{x} g(y) \, \mathrm{d}y$ is differentiable everywhere and $G'(x) = g(x)$ for all $x \in \mathbb{R}$. Indeed, $f + G$ fulfills all the requirement but is surely of unbounded variation.

Exercise 3.11' Define f to be the same function as is in Exercise 3.11. Prove that f is absolute continuous iff f is of bounded variation under the premise of $a, b > 0$.

Substitute the sine function in f with other periodic functions. Which properties can you specify for the same conclusion to hold? [This one is rather open and its answer is omitted here.]

Solution Only sufficiency is of interest. Note that it has to rely on the specific form of f. Suppose $a > b > 0$, since f' exists everywhere on $[c, 1]$ and is bounded whenever $c > 0$, f is sure to be absolutely continuous on $[c, 1]$.

The last piece of puzzle is filled by the continuity of $L(0, x)$. Given $\varepsilon > 0$, we can select a c such that $L(0, c) < \varepsilon$. For a finite collection of disjoint intervals (a_i, b_i) on $[0, 1]$, if none of them contains c we can split it into two parts:

$$\begin{aligned}\sum_{i=1}^{N} |F(b_k) - F(a_k)| &= \sum_{a_k \ge c} |F(b_k) - F(a_k)| + \sum_{b_k \le c} |F(b_k) - F(a_k)| \\ &\le \varepsilon + L(0, c) \le 2\varepsilon,\end{aligned}$$

as long as $\sum_{i=1}^{N} (b_k - a_k)$ is sufficiently small. Otherwise, split the only interval (a_k, b_k) that covers c into (a_k, c) and (c, b_k), and the variation only becomes greater. ■

Exercise 3.13 Show that the Cantor-Lebesgue function is not absolutely continuous.

Solution Note that $[0, 1] - \mathcal{C}$ is open. For arbitrary $\delta \in (0, 1)$, choose a finite number of intervals in $[0, 1] - \mathcal{C}$ whose length is at least $1 - \delta$ in total. Its complement in $[0, 1]$ is also a finite union of disjoint intervals, and the variation on these intervals is exactly 1. ■

Exercise 3.16 (TV Upper Bound) Show that if F is of bounded variation in $[a, b]$, then:

(a) $\int_a^b |F'(x)|\ \mathrm{d}x \leq T_F(a, b)$.

(b) The identity holds iff F is absolutely continuous.

As a result of (b), the formula $L = \int_a^b |z'(t)|\ \mathrm{d}t$ for the length of a rectifiable curve parametrized by z holds if and only if z is absolutely continuous.

Solution

(a) Clearly $|F'|$ is integrable. We begin by real-valued F. We write $F = F^+ - F^-$ as the difference of positive variations and negative variations. Fatou's lemma helps to establish the inequality on F^+ and F^-. Combined them together, we have

$$\begin{aligned}\int_a^b |F'(x)|\ \mathrm{d}x &= \int_a^b |F'^+(x) - F'^-(x)|\ \mathrm{d}x \\ &\leq \int_a^b F'^+(x)\ \mathrm{d}x + \int_a^b F'^-(x)\ \mathrm{d}x \\ &\leq T_{F^+}(a, b) + T_{F^-}(a, b) = T_F(a, b).\end{aligned}$$

Now we return to the general case where F is complex-valued. Let g be a step function such that $\|F' - g\|_{L_1} \leq \varepsilon$. Suppose G induces the partition $a = t_0 < \cdots < t_N = b$, we have

$$\begin{aligned}\int_a^b |F'(x)|\ \mathrm{d}x &\leq \varepsilon + \int_a^b |g(x)|\ \mathrm{d}x \\ &= \varepsilon + \sum_{i=1}^N \left| \int_{t_{i-1}}^{t_i} g(x)\ \mathrm{d}x \right| \\ &\leq 2\varepsilon + \sum_{i=1}^N \left| \int_{t_{i-1}}^{t_i} F'(x)\ \mathrm{d}x \right| \\ &\leq 2\varepsilon + \sum_{i=1}^N T_F(t_{i-1}, t_i) = 2\varepsilon + T_F(a, b).\end{aligned}$$

The last inequality needs further explanation: Denote the total variation of the real part and the imaginary part as $T_{\mathrm{Re}\, F}$ and $T_{\mathrm{Im}\, F}$ respectively. By invoking Minkowski's inequality, one can verify that

$$T_F(a, b) \geq \sqrt{T^2_{\mathrm{Re}\, F}(a, b) + T^2_{\mathrm{Im}\, F}(a, b)}. \tag{3.2}$$

Hence $\left| \int_{t_{i-1}}^{t_i} F'(x)\ \mathrm{d}x \right| \leq \sqrt{T^2_{\mathrm{Re}\, F}(t_{i-1}, t_i) + T^2_{\mathrm{Im}\, F}(t_{i-1}, t_i)} \leq T_F(t_{i-1}, t_i)$.

(b) Once the identity hold on $[a, b]$, it holds on any sub-interval. When F is real-valued and increasing, since $F(y) = \int_a^y F'(x)\ \mathrm{d}x + F(a)$ it must be absolutely continuous. Now suppose F is real-valued. It immediately follows that the last inequality is equal, which is to say F^+ and F^- are absolutely continuous. Hence, F is absolutely continuous

Finally we consider the general case. This time, $T_F(a, x)$ is absolutely continuous, which implies that $T_{\mathrm{Re}\, F}(a, x)$ and $T_{\mathrm{Im}\, F}(a, x)$ are both absolutely continuous. It then follows that F is absolutely continuous.

■

Exercise 3.17 Prove that if $\{K_\delta\}_{\delta>0}$ is a family of approximations to the identity, then

$$\sup_{\delta>0} |(f * K_\delta)(x)| \le cf^*(x)$$

for some constant $c > 0$ and all integrable f.

Solution We simply repeat here the intuition when bounding $|(f * K_\delta)(x) - f(x)|$.

$$\begin{aligned}
|(f * K_\delta)(x)| &\le \int_{\mathbb{R}^d} |f(x-y)| \cdot |K_\delta(y)| \, \mathrm{d}y \\
&\le A\delta^{-d} \int_{|y|\le\delta} |f(x-y)| \cdot |K_\delta(y)| \, \mathrm{d}y \\
&\qquad + A\delta \sum_{i=1}^{\infty} \frac{1}{2^{(i-1)(d+1)}\delta^{d+1}} \int_{2^{i-1}\delta\le|y|\le 2^i\delta} |f(x-y)| \, \mathrm{d}y \\
&\le A' f^*(x) + A'' \sum_{i=1}^{\infty} \frac{1}{2^{i-1}} f^*(x) \le cf^*(x).
\end{aligned}$$

Note that c is a constant dependent merely on d and A. ■

Exercise 3.19' (Equivalent Definition of Absolute Continuity) Suppose f is defined on $[a, b] \subseteq \mathbb{R}$. f is absolutely continuous on $[a, b]$ iff given $\varepsilon > 0$ there exists a $\delta > 0$ such that for any open set $E \subseteq [a, b]$ with $m(E) < \delta$ we have

$$\sum_{k=1}^{\infty} |f(a_k) - f(b_k)| \le \varepsilon,$$

where $E = \bigcup_{k=1}^{\infty}(a_k, b_k)$ is the unique decomposition of E.

It immediately follows that E can be extended to measurable sets in lieu of open sets.

Solution A proof of the necessity side is sufficient here. Such a δ exists when it comes to a finite union. We will show that it works for a countable union as well.

Let $m(E) < \delta$. For any given N, $E_N = \cup_{k=1}^{N}(a_k, b_k)$ is a finite union of length less than δ, and we deduce that $\sum_{i=1}^{N} |f(a_k) - f(b_k)| \le \varepsilon$. Hence the series converges to a point no greater than ε. ■

Exercise 3.19 Show that if $f : \mathbb{R} \to \mathbb{R}$ is absolutely continuous, then f maps sets of measure zero to sets of measure zero. What directly follows is that f maps measurable sets to measurable sets.

Hint Invoke Exercise 3.19'. For an interval (a_k, b_k) it will be helpful to interpose the points $\arg\max_{x\in[a_k,b_k]} f(x)$ and $\arg\min_{x\in[a_k,b_k]} f(x)$. If there are multiple maximum (or minimum) points, interpose any one of them.

Exercise 3.20' Suppose f is an increasing absolutely continuous function defined on $[a, b]$. Let $A = F(a)$ and $B = F(b)$. Prove that $m(\mathcal{O}) = \int_{F^{-1}(\mathcal{O})} F'(x) \, \mathrm{d}x$ for any open set $\mathcal{O} \subseteq [A, B]$.

Further conclude that for a set $\mathcal{E}$ of measure zero the set $F^{-1}(\mathcal{E}) \cap \{F'(x) > 0\}$ has measure zero.

Solution Decomposite $\mathcal{O}$ as union of disjoint intervals (A_k, B_k) in $[A, B]$. Let $F^{-1}(A_k) = a_k$ and $F^{-1}(B_k) =$

b_k. Since f is absolutely continuous we have

$$\int_{a_k}^{b_k} F'(x)\,\mathrm{d}x = F(b_k) - F(a_k) = m((A_k, B_k)).$$

It then follows that

$$\begin{aligned} m(\mathcal{O}) = \sum_{k=1}^{\infty} m((A_k, B_k)) &= \sum_{k=1}^{\infty} \int_{a_k}^{b_k} F'(x)\,\mathrm{d}x \\ &= \int_{\cup_{k=1}^{\infty} F^{-1}((A_k, B_k))} F'(x)\,\mathrm{d}x \\ &= \int_{F^{-1}(\cup_{k=1}^{\infty}(A_k, B_k))} F'(x)\,\mathrm{d}x \\ &= \int_{F^{-1}(\mathcal{O})} F'(x)\,\mathrm{d}x. \end{aligned}$$

Next we select a sequence of open sets $\mathcal{O}_i \supseteq \mathcal{E}$ such that $m(\mathcal{O}) \leq 1/i$. Let $\mathcal{E}_n = \cap_{i\leq n}\mathcal{O}_i$ and $\mathcal{E}_n \searrow \mathcal{E}'$. Clearly we have $m(\mathcal{E}_n) \searrow m(\mathcal{E}')$. Moreover, since $F'(x)\,\chi_{F^{-1}(\mathcal{E}_n)} \searrow F'(x)\,\chi_{F^{-1}(\mathcal{E}')}$, by invoking dominated convergence theorem we have

$$\int_{F^{-1}(\mathcal{E}_n)} F'(x)\,\mathrm{d}x \to \int_{F^{-1}(\mathcal{E}')} F'(x)\,\mathrm{d}x.$$

Hence $0 = m(\mathcal{E}') = \int_{F^{-1}(\mathcal{E}')} F'(x)\,\mathrm{d}x = \int_{F^{-1}(\mathcal{E}_n)\cap\{F'(x)>0\}} F'(x)\,\mathrm{d}x$, which implies that it is a.e. zero. Equivalently, it is to say that $F^{-1}(\mathcal{E}_n) \cap \{F'(x) > 0\}$ has measure zero. What remains is to notice $F^{-1}(\mathcal{E}) \subseteq F^{-1}(\mathcal{E}')$. ■

Exercise 3.20 The assumption on the function f is the same as in Exercise 3.20'. Let E denote a measurable subset of $[A, B]$. Show that the set $F^{-1}(E) \cap F'(x) > 0$ is measurable.

It is worth pointing out that $F^{-1}(E)$ itself may not be measurable.

Hint E can be expressed as the difference of a G_δ and a set of measure zero.

Here is an example for $F^{-1}(E)$ to be not measurable. Take a Cantor-like set $\mathcal{C}$ on $[0, 1]$ with positive measure. Let $F(x) = m([0, x]/\mathcal{C})$ for all $x \in [0, 1]$. Of course, F is 1-Lipschitz and is thus absolutely continuous. Moreover, F is strictly increasing. What makes F more interesting is as follows: Clearly $F(\mathcal{C}) = F([0, 1]) - F([0, 1]/\mathcal{C})$. Since $[0, 1]/\mathcal{C}$ is open and on each interval contained in it F perserves its measure, $m([0, 1]/\mathcal{C}) = m(F([0, 1]/\mathcal{C}))$. There we have $m(F(\mathcal{C})) = m(F([0, 1])) - m([0, 1]/\mathcal{C}) = 0$. Choose a subset $K \subseteq E$ that is not measureble. $F(K)$ has measure 0, but $F^{-1}(F(K)) = K$.

Exercise 3.21 (The Change of Variable Formula) Let F be absolutely continuous and increasing on $[a, b]$ with $F(a) = A$ and $F(b) = B$. Suppose f is any measurable function on $[A, B]$.

(a) Show that $f(F(x))F'(x)$ is measurable on $[a, b]$. Note, however, that $f(F(x))$ need not to be measurable.

(b) Prove that change of variable formula: If f is integrable on $[A, B]$, then so is $f(F(x))F'(x)$, and

$$\int_A^B f(y)\,\mathrm{d}y = \int_a^b f(F(y))F'(y)\,\mathrm{d}y. \tag{3.3}$$

Solution It begins by noting that $S_{x,y} = \{f(F(x)) > x \wedge F'(x) > y\}$ is a measurable set when $x \in \mathbb{R}$ and

$y > 0$. So does $T_{x,y} = \{f(F(x)) < x \wedge F'(x) > y\}$. Lastly, the inverse image of 0 can be expressed as $\{x : f(F(x)) = 0 \wedge F'(x) > 0\} \cup \{x : F'(x) = 0\}$, which is also measurable.

To establish the identity, it may be tempting to approximate f with step functions. But you will soon get trapped in bounding $\|f(F(y))F'(y) - g(F(y))F'(y)\|_{L^1([a,b])}$. Here we take alternative steps starting from the case where f is non-negative.

First we look into $f(y) = \chi_E$, where E is a measurable set in $[A, B]$. From Exercise 3.20', we have

$$\begin{aligned} m(E) &= \int_{F^{-1}(E)\cap\{F'(x)>0\}} F'(x)\, \mathrm{d}x \\ &= \int_a^b F'(x)\, \chi_{F^{-1}(E)\cap\{F'(x)>0\}}\, \mathrm{d}x \\ &\leq \int_a^b f(F(y))F'(y)\, \mathrm{d}y. \end{aligned}$$

On the other hand,

$$\begin{aligned} m(E) &= \int_{(F^{-1}(E)\cap\{F'(x)>0\})\cup\{F'(x)=0\}} F'(x)\, \mathrm{d}x \\ &\geq \int_a^b f(F(y))F'(y)\, \mathrm{d}y. \end{aligned}$$

Thus, $f = \chi_E$ satisfies Equation 3.3. Immediately simple functions satisfies the equation as well. The last step is to note that there is a sequence of simple functions $\{f_n\}$ which pointwise increases and converges to f everywhere on $[A, B]$. By a limiting argument on both sides of the identity, f satisfies this formula. As the last step, write $f = f^+ - f^-$. According to the additivity on both sides, the change of variable formula is valid for all integrable f. ■

Exercise 3.23 Let F be continuous on $[a, b]$ and suppose $(D^+F)(x) \geq 0$ for every $x \in [a, b]$. Then F is increasing on $[a, b]$.

Solution

■

Exercise 3.24 (Lebesgue Decomposition) Suppose F is an increasing function on $[a, b]$ of finite value.

(a) Prove that we can write

$$F = F_A + F_C + F_J,$$

where each of the functions F_A, F_C and F_J is increasing and:

1. F_A is absolutely continuous.
2. F_C is continuous, but $F_C'(x) = 0$ for a.e. x.
3. F_J is a jump function.

(b) Moreover, each component is uniquely determined up to an additive constant.

(c) Show a similar conclusion for F of bounded variation.

You may recall **CDF** (Cumulative Density Function) in probability theory. In it continuous distribution corresponds to the F_A component and discrete distribution the F_J component. As you may guess, singular distribution interprets exactly the F_C term.

Solution After knowing all those theorems, the decomposition is readily within our reach. Suppose F is increasing, and denote the set of points at which F discontinue as $\{x_n\}_{n=1}^{\infty}$. Let $a_n = D_+(x_n) - D^-(x_n)$ be the gap at x_n. As usual, if we let

$$j_n(x) = \begin{cases} 0 & \text{if } x < x_n, \\ \frac{F(x_n) - D^-(x_n)}{a_n} & \text{if } x = x_n, \\ 1 & \text{if } x > x_n, \end{cases}$$

then we define F_J by

$$F_J(x) = \sum_{i=1}^{\infty} a_n j_n(x).$$

It follows immediately that $G := F - F_J$ is increasing and continuous. Since G' exists a.e. and is integrable, $F_A(x) = \int_a^x G'(y)\ \mathrm{d}y$ is absolutely continuous. Now look into the differentiation. G' is identical to both F' and F'_A, where the former arises from $F'_J = 0$ a.e. and the latter arises from Lebesgue differentiation theorem. Thus, defined by $G - F_A$, F_C is continuous but F'_C vanishes a.e.. A similar argument holds for the two variations $P_F(a, x)$ and $N_F(a, x)$, so it is valid for F of bounded variation.

We shall next establish the uniqueness of this decomposition. At this point, it is convenient to view each component as the representative of equivalent classes formed by constant drift. Suppose $F = F_A + F_C + F_J = \overline{F}_A + \overline{F}_C + \overline{F}_J$. The two jump functions F_J and $\overline{F}_J$ must cancel out at each point of its discontinuities to induce a continuous function, so they are the same [which should be formally expressed in analytic definition of F_J and $\overline{F}_J$, but is omitted here to remove clutter]. Next we claim that F'_A and $\overline{F}'_A$ agree a.e. due to the fact that $F'_C - \overline{F}'_C$ vanishes a.e.. This property, along with absolute continuity, implies that F_A and $\overline{F}_A$ are identical.

The uniqueness does not rely on the monotonicity of F, so the argument above still applies to F of bounded variation. ■

Exercise 3.25 The following shows the necessity of allowing for general exceptional sets of measure zero in the differentiation theorems. Let E be any set of measure zero in $\mathbb{R}^d$. Show that:

(a) There exists a non-negative integrable f in $\mathbb{R}^d$, so that

$$\liminf_{\substack{m(B) \to 0 \\ x \in B}} \frac{1}{m(B)} \int_B f(y)\ \mathrm{d}y = \infty.$$

(b)

Solution ■

Exercise 3.26 The Vitali covering lemma introduced in this chapter makes the argument that Lebesgue measure remains the same if we are to cover a set with balls much easier to prove. Since a ball is a union of almost disjoint cubes, $m_*(E) \le m_*^{\mathcal{B}}(E)$, where E is any set and $m_*^{\mathcal{B}}$ denotes the new Lebesgue measure we consider here. Your task is show that the reverse order is true as well, i.e., $m_*(E) \ge m_*^{\mathcal{B}}(E)$.

Solution It is sufficient to prove that an open cube C of side length l has a mesure of l^d under $m_*^{\mathcal{B}}$. For any given $\varepsilon > 0$, we construct a covering of balls with $\sum_{i=1}^{\infty} m_*^{\mathcal{B}}(B_i) \le l^d + \varepsilon$.

First, we argue that a set E of measure zero also has measure zero under $m_*^{\mathcal{B}}$. Indeed, by choosing a cube covering of E and replacing each cube with the smallest ball that contains it, the sum of volume inflates by a constant factor depending merely on d.

Given $x \in C$, there exists a ball $B(x, r_x) \subseteq C$ centered at x. Let $\mathcal{B} = \{B(x,r) : x \in C, 0 \le r \le r_x\}$ be a Vitali covering of C. Run the following procedure indefinitely.

Procedure:

1. Let $C = C$, $\mathcal{B} = \mathcal{B}$, $\mathcal{B}' = \emptyset$;
2. Iterate the following indefinitely:
   (a). Choose a compact set $K \subseteq C$ with $m(K) \ge m(C)/2$;
   (b). Since $\mathcal{B}$ is a covering of K, select a finite collection $\tilde{\mathcal{B}}$ that covers K;
   (c). From $\tilde{\mathcal{B}}$ select a disjoint sub-collection $\tilde{\mathcal{B}}'$ that covers K with $m(\bigcup_{B\in\tilde{\mathcal{B}}'} B) \ge m(K)/3^d$.
   (d). Let $\mathcal{B}' = \mathcal{B}' \cup \tilde{\mathcal{B}}'$;
   (e). Update $C = C / \bigcup_{B\in\tilde{\mathcal{B}}'} \overline{B}$ which is still open;
   (f). Update $\mathcal{B} = \{B \in \mathcal{B} : B \subseteq C\}$. Note that B is still a covering of the newly updated C.

Logically, in the end we get a denumerable collection $\mathcal{B}' = \{B_i\}_{i=1}^{\infty}$ such that $m(\overline{C} - \bigcup_{i=1}^{\infty} B_i) = 0$. To make it a valid covering, note that the only points $\mathcal{B}'$ fails to cover can be covered by balls with a total volume arbitrarily small, as is argued before. So far, the prove is completed and a direct spinoff is the rotation invariance of Lebesgue measure. ■

Exercise 3.29 Let $\Gamma = \{z(t), a \le t \le b\}$ be a curve, and suppose it satisfies a Lipschitz condition with exponent α, where $1/2 \le \alpha \le 1$, i.e.,

$$|z(t) - z(t')| \le A|t-t'|^{\alpha} \quad \text{for all } t, t' \in [a,b].$$

Show that $m(\Gamma^{\delta}) = O(\delta^{2-1/\alpha})$ for $0 < \delta \le 1$.

Solution A coarse estimation will be sufficient to give the desired order. Divide $[a,b]$ into $l := \dfrac{b-a}{\delta^{1/\alpha}} + 1$ intervals of equal length, which we call I_i. As a consequence, for each pair $t, t' \in I_i$ there is $|z(t) - z(t')| \le \delta$. Momentarily write $\Gamma_i = \{z(t), t \in I_i\}$ and denote the line segment connecting z-values of I_i's left boundary and right boundary as J_i. One can verify that

$$\Gamma^{\delta} \subseteq \bigcup_{k=1}^{l} \Gamma_k^{\delta} \quad \text{and} \quad \Gamma_i^{\delta} \subseteq J_i^{(A+1)\delta} \text{ for all } i.$$

We have

$$m(\Gamma^{\delta}) \le \sum_{i=1}^{l} m(J_i^{(A+1)\delta}) \le \sum_{i=1}^{l} 2(A+1)\delta \cdot (2A+3)\delta = O(\delta^{2-1/\alpha}).$$

■

Exercise 3.30 (Globally Bounded Variation) A bounded function F is said to be of bounded variation on $\mathbb{R}$ if F is of bounded variation on any finite sub-interval $[a,b]$, and $\sup_{a,b} T_F(a,b) < \infty$. Prove that such an F enjoys the following two properties:

(a) $\int_{\mathbb{R}} |F(x+h) - F(x)| \,\mathrm{d}x \le A|h|$ for some constant $A > 0$ and all $h \in \mathbb{R}$.

(b) $\int_{\mathbb{R}} |F(x)\varphi'(x)| \,\mathrm{d}x \le A$, where φ ranges over all C^1 functions of bounded support with $\sup_{\mathbb{R}} |\varphi| \le 1$.

Solution We may assume that $h > 0$. In the following, let $A = \sup_{a,b} T_F(a,b) < \infty$. First we claim that for any partition π such that it has at most finitely many partition points in each $[a,b]$ has a TV at most A. Using this fact, we may massage the integral as follows:

$$\begin{aligned}\int_{\mathbb{R}} |F(x+h) - F(x)| \, \mathrm{d}x &= \sum_{k=-\infty}^{\infty} \int_{(k-1)h}^{kh} |f(x+h) - f(x)| \, \mathrm{d}x \\ &= \int_0^h \sum_{k=-\infty}^{\infty} |f(x+kh) - f(x+(k-1)h)| \, \mathrm{d}x \\ &\leq \int_0^h A \, \mathrm{d}x = Ah.\end{aligned}$$

Now we come to (b). Suppose φ is supported on $[-C, C]$. Note that φ is Lipschitz by assumption that it is indeed C^1. By the formula of integral by parts,

$$\begin{aligned}\int_{\mathbb{R}} |F(x)\varphi'(x)| \, \mathrm{d}x &= \int_{-C}^{C} |F(x)\varphi'(x)| \, \mathrm{d}x \\ &= \left| F(x)\varphi(x)|_{-C}^{C} - \int_{-C}^{C} F'(x)\varphi(x) \, \mathrm{d}x \right| \\ &\leq \int_{-C}^{C} |F'(x)| \cdot |\varphi(x)| \, \mathrm{d}x \leq \int_{\mathbb{R}} |F'(x)| \, \mathrm{d}x.\end{aligned}$$

Let $F_n(x) := \dfrac{F(x+1/n) - F(x)}{1/n}$, then $F_n(x) \to F'(x)$ a.e. as $n \to \infty$. By invoking Fatou's lemma, we have

$$\int_{\mathbb{R}} |F'(x)| \, \mathrm{d}x \leq \liminf_{n\to\infty} \int_{\mathbb{R}} |F_n(x)| \, \mathrm{d}x \leq A.$$

Putting these inequalities altogether, we have reached the desired conclusion. ■

Exercise 3.31 For a function $f(x)$ defined on $[a, b]$, we consider the curve given by $x(t) = t, y(t) = f(t)$ where $t \in [a, b]$. Now we can talk about its length.

Show that the length of Cantor-Lebesgue function (of trinary $\mathcal{C}$) has length 2.

Solution In fact, define $L(u) = T_F(0, u)$ for $u \in [0, 1]$ to be the length of the segment $t \in [0, x]$. We have $L(u) = u + F(u)$.

To upper bound $L(u)$, we use the fact that

$$T_z(a, b) \leq T_{\mathrm{Re}\, z}(a, b) + T_{\mathrm{Im}\, z}(a, b) \tag{3.4}$$

for any curve z in $\mathbb{R}^2$. The notation here is the same as in euqation 3.2. Now that $T_x(0, u) = u$ and $T_y(0, u) = F(u)$, there is $L(u) \leq u + F(u)$.

To lower bound $L(u)$, we give a specific partition for any given $\varepsilon > 0$. Let $\mathcal{C}_k$ denote the Cantor set constructed after the k-th step. Choose k sufficiently large and we can make $m(\mathcal{C}_k) \leq \varepsilon$. We consider partition given by boundaries of $\mathcal{C}_k \cap [0, u]$. F is flat on its complement in $[0, u]$, and would enjoy a total variation of at least $u - \varepsilon$. On the rest segment where f is not flat, TV is no less than $T_y(0, u) = F(u)$. Hence $L(u) \geq u - \varepsilon + F(u)$. Since ε can be arbitrarily close to 0, the claim has been proved. ■

3.2 Problems

Problem 3.1 Prove the following variant of the **Vitali covering lemma**: If E is covered in the Vitali sense by a family $\mathcal{B}$ of balls, and $0 < m_*(E) < \infty$, then for every $\eta > 0$ there exists a disjoint collection of balls $\{B_j\}_{j=1}^{\infty}$ in $\mathcal{B}$ such that

$$m_*\left(E/\bigcup_{j=1}^{\infty} B_j\right) = 0 \quad \text{and} \quad \sum_{j=1}^{\infty} |B_j| \leq (1+\eta)m_*(E).$$

Solution It suffices to prove for the case where E is an open set and $\eta = 0$. The reason is as follows: Select an open set $\mathcal{O}$ containing E with $m(\mathcal{O}) \leq (1+\eta)m_*(E)$. Keep those balls in $\mathcal{B}$ which is fully contained in $\mathcal{O}$, and we get a new Vitali family $\mathcal{B}'$ of balls. From here onward, we view $E' = \bigcup_{B\in\mathcal{B}'} B$, which clearly is an open set, as our original E. The edge of this massage is that the last condition is automatically satisfied.

What follows is a vanilla iterated invocation of Vitali covering lemma. Suppose now we are to fill in an open set E. Choose a compact set $\mathcal{E} \subseteq E$ with $m(\mathcal{E}) \geq m(E)/2$. Then there is a finite covering $\{B_{1,i}\}_{i=1}^{N_1}$ of $\mathcal{E}$. From this covering we may select a disjoint sub-collection $\{B_{1,i}\}_{i=1}^{M_1}$ (after re-indexing) whose coverage is more than $3^{-d}/2$. Let $E_2 = E - \bigcup_{i=1}^{M_1} \overline{B}_{1,i}$ and $\mathcal{B}_2$ to be the collection of balls in $\mathcal{B}$ which does not intersect with $\bigcup_{i=1}^{M_1} \overline{B}_{1,i}$. Continue the above process indefinitely. The collection $\bigcup_{j=1}^{\infty}\bigcup_{i=1}^{M_j} B_{j,i}$ is indeed the desired sub-collection since

$$m\left(E/\bigcup_{j=1}^{N}\bigcup_{i=1}^{M_j} B_{j,i}\right) \leq (1-\frac{3^{-d}}{2})^N.$$

■

Problem 3.2 (1d Covering Lemma) Suppose $\{I_j\}_{j=1}^{N}$ gives a finite collection of open intervals in $\mathbb{R}$. Then there are two finite sub-collections $\{I_k^1\}_{k=1}^{K}$ and $\{I_l^2\}_{l=1}^{L}$, so that each sub-collection consists of mutually disjoint intervals and

$$\bigcup_{j=1}^{N} I_j = \bigcup_{k=1}^{K} I_k^1 \cup \bigcup_{l=1}^{L} I_l^2.$$

Actually, $1/2$ is the best possible coefficient in lemma 1.2 [in the book] when $d = 1$. I.e., we can always select a disjoint sub-collection of intervals from a finite collection of them such that at least $1/2$ of the area is now covered.

Solution It in fact can be solved by the following greedy algorithm in **P** time. One thing to mention is that we may assume I_j to be finite. Otherwise, substitute ∞ with any real numbers that is sufficiently far from 0. We shall omit the discussion of its validity and time efficiency here. The discussion per se will substantiate the theorem at the same time. ■

Algorithm 1: 1d Covering Problem

Input: N open intervals of finite measure $I_j = (a_i, b_i)$

Output: Sets of indices, indicating the two sub-collections

while *there is an interval I covered by union of others* **do**
  Remove I;
end
Sort the rest intervals according to a_i in an increasing order;
Assume N' intervals are left. During sorting, maintain a mapping $f : [N'] \to [N]$ to restore the original index;
Initialize $C_1 = \emptyset$, $C_2 = \emptyset$;
for i *in* $[N']$ **do**
  Insert $f(i)$ into $C_{i \bmod 2}$;
end
return C_1, C_2

Problem 3.3 There is no direct analogue of Problem 3.2 in higher dimensions. However, a full covering is afforded by the Besicovitch covering lemma. A version of this lemma states that there is an integer $N = N(d)$ (i.e., only dependent on d) with the following property. Suppose E is any bounded set in $\mathbb{R}^d$ that is covered by a collection $\mathcal{B}$ of balls in the (strong) sense that for each $x \in E$, there is a $B \in \mathcal{B}$ whose center is x. Then, there are N sub-collections $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_N$, and moreover,

$$E \subseteq \bigcup_{B \in \mathcal{B}'} B, \quad \text{where } \mathcal{B}' = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_N.$$

Problem 3.5 Suppose that F is continuous on $[a, b]$, $F'(x)$ exists for *every* $x \in [a, b]$, and $F'(x)$ is integrable. Then F is absolutely continuous and

$$F(b) - F(a) = \int_a^b F'(x)\, \mathrm{d}x.$$

Remark This problem it quite subtle. As is emphasized in italic, "everywhere" matters a lot. However, I did not see the meaning of the hint in the textbook. If you do have any idea, please contact me.

Lemma 3.1

Suppose f is continuous on $[a, b]$, E is a measurable subset of (a, b) and f is differentiable on every point of E. If

$$\sup_{x \in E} |f'(x)| \leq M < \infty,$$

then $f(E)$ is measurable. Moreover,

$$m(f(E)) \leq M \cdot m(E).$$

♡

Proof We first claim the result for E with measure zero. Given $\varepsilon > 0$, define

$$E_{1/n} = \{x \in E : |f(x) - f(y)| < (M + \varepsilon)|x - y| \text{ for all } |x - y| \leq 1/n\}.$$

Clearly, $E_{1/n} \nearrow E$, and $f(E_{1/n}) \nearrow f(E)$. It suffices to show that $f(E_{1/n})$ has measure zero.

Since $m(E_{1/n})$ has measure zero, we can cover it with an open set $\mathcal{O}$ of measure δ. Let $\mathcal{O} = \bigcup_{i\in\mathbb{N}} I_i$ be the unqiue decomposition and divide each I_i into sub-intervals $I_{i,j}$ whose length are less than $1/n$. Now we have

$$\begin{aligned} m_*(f(E_{1/n})) &\leq \sum_{\exists\, x\in E_{1/n} \text{ s.t. } x\in I_{i,j}} m(f(I_{i,j})) \\ &\leq \sum_{\exists\, x\in E_{1/n} \text{ s.t. } x\in I_{i,j}} M\cdot 2m(I_{i,j}) \\ &\leq 2M\sum_{i,j} m(I_{i,j}) \leq 2M\delta. \end{aligned}$$

Note that δ can be arbitrarily small, we have $m_*(f(E_{1/n})) = 0$.

Next, we prove it for every closed set. It is indeed compact in this case, and therefore, $f(E)$ is also compact. Select an open set $\mathcal{O} \subseteq E$ s.t. $m(\mathcal{O} - E) \leq \varepsilon$, and we construct a covering of E by adding to it $I_x = (a_x, b_x)$ for each $x \in E$ with the property that $|f(y) - f(x)| \leq (M+\varepsilon)|y-x|$ for any $y \in I_x$. We then have a finite sub-collection $\{I_i\}_{i=1}^N$ of E. By appropriately resetting its boundary, we can ensure that $x_j \in \overline{I_j}$, $\overline{I_j}$ does not overlap with others except on the boundary and that $\overline{I_i}$ covers the same area. This way, on each I_j the image of f has measure less than $M\cdot m(I_j)$, and it follows that $m(f(E)) \leq M(m(E)+\varepsilon)$. Since ε is arbitrary, $m(f(E)) \leq M\cdot m(E)$.

As the last step, note every measurable set E is a (disjoint) union of F_σ and a set of measure zero. Let $E = F \cup \bigcup_n \mathcal{E}_n$, where $m(F) = 0$ and $\mathcal{E}_n$ is closed and converges to E.

$$m(f(E)) = m(f(F)) + \lim_{k\to\infty} m(f(\bigcup_{n\leq k}\mathcal{E}_n)) \leq \lim_{k\to\infty} M\cdot m(\bigcup_{n\leq k}\mathcal{E}_n) \leq M\cdot m(E).$$

□

Lemma 3.2

Suppose f is continuous on $[a,b]$ and f' exists for every $x\in(a,b)$ and f' is integrable. Then

$$m(f(a,b)) \leq \int_a^b |f'(x)|\, dx.$$

♡

Proof Given $\varepsilon > 0$, let $E_k := \{x : k\varepsilon \leq |f'(x)| < (k+1)\varepsilon\}$. We have

$$\begin{aligned} \int_a^b |f'(x)|\, \mathrm{d}x &\geq \sum_{k=0}^{\infty}(k+1)\varepsilon\cdot m(E_k) - (b-a)\varepsilon \\ &\geq \sum_{k=0}^{\infty} m(f(E_k)) - (b-a)\varepsilon = m(f(a,b)) - (b-a)\varepsilon. \end{aligned}$$

Choosing ε sufficiently close to 0, we deduce the inequality. □

Solution Back to the original problem.

We are now in a position ready to show that F is absolutely continuous. Note that $G(x) = \int_a^x |F'(y)|\, \mathrm{d}y$ is absolutely continuous. For $\varepsilon > 0$, we choose the corresponding δ in the definition of absolute continuity and we claim any disjoint collection $\bigcup_{i=1}^N [a_i, b_i] \subseteq [a,b]$ with $\sum_{i=1}^N |b_i - a_i| \leq \delta$, there is

$$\sum_{i=1}^N |F(b_i) - F(a_i)| \leq \sum_{i=1}^N m(F(a_i,b_i)) \leq \sum_{i=1}^N \int_{a_i}^{b_i} |F'(x)|\, \mathrm{d}x \leq \varepsilon.$$

Upon the fact that F is absolutely continuous, the equation that $F(b) - F(a) = \int_a^b F'(x)\, \mathrm{d}x$ holds. ■

Problem 3.7 (Continuous Everywhere but Differentiable Nowhere Function; Fractal Curve; Weierstrass Function)

(a) We first present a variant of Weierstrass function known as Weierstrass sawtooth function. Let $f(x) = 1 - |2\{x/2\} - 1|$ for $x \in \mathbb{R}$, where $\{x\}$ denotes the fraction part of x. Construct the function

$$F(x) = \sum_{n=0}^{\infty} f_n(x) = \sum_{n=0}^{\infty} 2^{-n} f(2^n x)$$

defined on $\mathbb{R}$. Show that F is continuous everywhere but differentiable nowhere. An illustration of this property is shown below.

(b) Consider the function

$$g_1(x) = \sum_{n=0}^{\infty} 2^{-n} \exp(2\pi i 2^n x).$$

1. Prove that g_1 satisfies $|g_1(x) - g_1(y)| \le A_\alpha |x - y|^\alpha$ for each $0 < \alpha < 1$.
2. However, g_1 is nowhere differentiable, hence not of bounded variation.

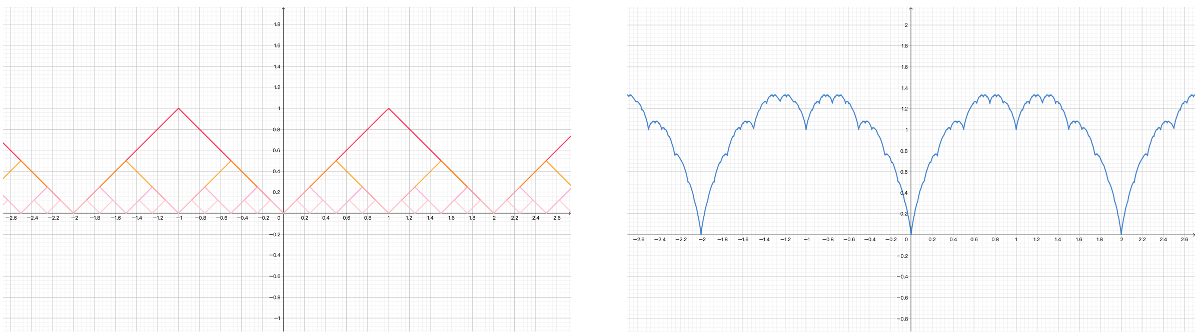

Figure 3.1: Illustration of WSF. The diagram on the left depicts the first four terms in the summation, while the one to its right delineates the $F(x)$ itself. It gives you a sense of how bumpy $F(x)$ is.

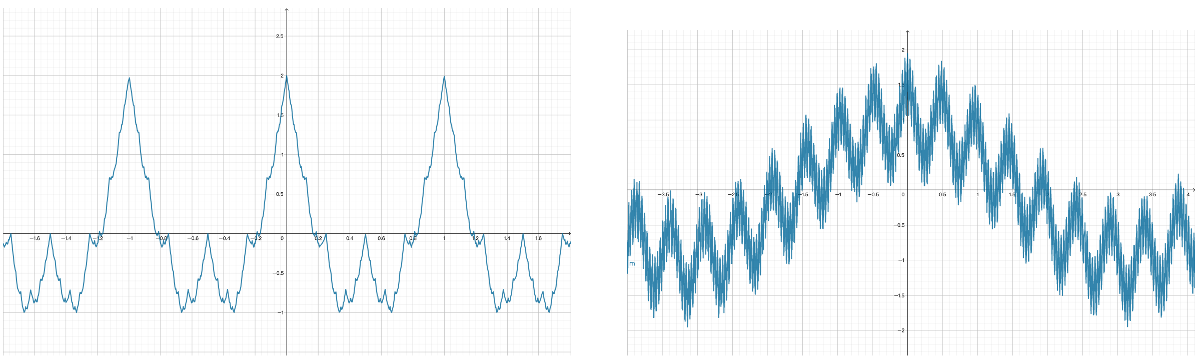

Figure 3.2: The left diagram visualizes the real part of the function in (b). The right one corresponds to Weierstrass function with $f(x) = \sum_{i=0}^{\infty} 2^{-n} \cos(13^n \pi x)$.

Solution The contintuity of F and g_1 are obvious due to uniform convergence of the series in the summation. To show that they are nowhere differentiable, we approach it by claiming that the upper and lower limit does not agree as it goes to x.

(a) For a given $x \in \mathbb{R}$, there is an N so that $f_n(x) \neq 2^{-n}$ for any $n \ge N$. We consider the partial sum formed by $F_n(x) = \sum_{i=0}^{n} f_i(x)$ when $n \ge N$. Let $x_0 = x$ be sandwiched between $a = \frac{m}{2^{n-1}}$ and $b = \frac{m+1}{2^{n-1}}$. Note that x_0 can be equal to a or b but the middle point $c = \frac{2m+1}{2^n}$. Let x_2 be the mirror of x_0 w.r.t. c

and $x_1 \in [a,b]$ be the only point s.t. $\min(|a-x_0|,|b-x_0|) = |c-x_1|$ and $(x_0-c)(x_1-c) \geq 0$. By assumption, x_0, x_1, x_2 are distinct.

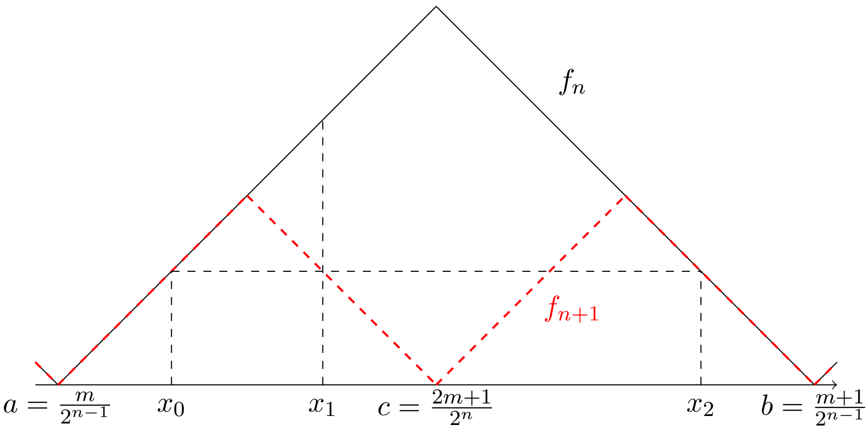

Figure 3.3: Geometric meaning of x_0, x_1 and x_2

Figure 3.3 illustrates this idea, but it only dislays the situation where x_0 resides in the leftmost quarter. We claim that the slope at x_0 with respect to x_1 and x_2 differs exactly by 1. Hence,

$$\limsup_{y\to 0}\left|\frac{F(y)-F(x)}{y-x}\right| - \liminf_{y\to 0}\left|\frac{F(y)-F(x)}{y-x}\right| \geq \limsup_{n\to\infty}\left|\frac{F(x_1)-F(x_0)}{x_1-x_0} - \frac{F(x_2)-F(x_0)}{x_2-x_0}\right| = 1.$$

A couple of observations come in order. First, $f_k(x_0) = f_k(x_1) = f_k(x_2)$ when $k > n$. Second, $\frac{f_l(x_1)-f_l(x_0)}{x_1-x_0}$ equals $\frac{f_l(x_2)-f_l(x_0)}{x_2-x_0}$ when $l < n$. Then the only difference roots in f_n term. While $f_n(x_0) = f_n(x_2)$ always, the slop between x_0 and x_1 will be either 1 or -1, depending on whether x_0 is to the left of c.

(b.1) Denote $\Delta = |x-y|$ for convenience. Given α, for any $\Delta > 0$, we must have

$$\begin{aligned}
|g_1(x)-g_1(y)| &\leq \sum_{i=0}^{\infty} 2^{-n}|\exp(2\pi i 2^n x) - \exp(2\pi i 2^n y)| && \triangleright \text{ Triangle inequality}\\
&= \sum_{i=0}^{\infty} 2^{-n}|\exp(2\pi i 2^n \Delta) - 1|\\
&\leq 2^{1-\alpha}\sum_{i=0}^{\infty} 2^{-n}|\exp(2\pi i 2^n \Delta) - 1|^{\alpha} && \triangleright |\exp(2\pi i 2^n \Delta) - 1| \leq 2\\
&\leq 2^{1-\alpha}\sum_{i=0}^{\infty} 2^{-n}\cdot|2\pi i 2^n \exp(2\pi i 2^n \Delta')\Delta|^{\alpha}\\
& && \triangleright \text{ Mean value theorem, where } \Delta' \in (0,\Delta)\\
&= 2\pi^{\alpha}\sum_{i=0}^{\infty} 2^{-(1-\alpha)n}\Delta^{\alpha} = A_{\alpha}\Delta^{\alpha}.
\end{aligned}$$

This yields the desired result.

(b.2)

■

Problem 3.9 Let $\mathcal{R}$ denote the set of all rectangles in $\mathbb{R}^2$ which needs *not* to contain the origin, and with sides parallel to the coordination axis. Consider the maximal operator associated to this family, namely

$$f_{\mathcal{R}}^*(x) = \sup_{R \in \mathcal{R}} \frac{1}{m(R)} \int_R |f(x-y)| \, \mathrm{d}y.$$

(a) Then, $f \mapsto f_{\mathcal{R}}^*$ does not satisfy the weak type inequality

$$m(\{x : f_{\mathcal{R}}^*(x) > \alpha\}) \leq \frac{A}{\alpha} \|f_1\|_{L^1}$$

for all $\alpha > 0$, all integrable f, and some $A > 0$.

(b) Using this, show that there exists $f \in L^1(\mathbb{R}^2)$ so that for $R \in \mathcal{R}$

$$\limsup_{\operatorname{diam}(R) \to 0} \frac{1}{m(R)} \int_{\mathbb{R}} f(x-y) \, \mathrm{d}y = \infty \quad \text{for almost every } x.$$

Here $\operatorname{diam}(R) = \sup_{x,y \in R} |x-y|$ equals the diameter of the rectangle.

Solution If we constrain $\mathcal{R}$ to be the collection of all cubes, we will get back to the original weak type inequality. The intuition lies in that if we can make f quite large near in the vicinity of axis, then an ideal candidate to near maximize the quantity will hopefully contain the axis.

It can be verified that $f(x,y)$ defined as follows is valid for the claim.

$$f(x,y) = \begin{cases} \frac{1}{\sqrt{|xy|}} & \text{if } 0 < |x|, |y| \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

■

3.3 Supplementaries

Chapter 4 Hilbert Spaces: An Introduction

Key words

- *Examples of Hilbert Spaces:* $L^2(E)$, $\ell^2(\mathbb{Z})$.
- *Completeness & Separability*
- *Orthonormal basis*
- *Completion of a pre-Hilbert space*
- *Linear transformations:*
  - *Unitary mappings*
  - *Orthogonal projections*
  - *Linear functionals*
  - *Compact operators*
  - *Integral operators*
- *Adjoints*
- *Operator norm*
- *Hardy space* $H^2(\mathbb{D})$
- *Hilbert-Schmidt operators*
- *Spectral theorem*
- *Diagonalization*

Unless otherwise noted, the Hilbert spaces we deal with in this chapter are all assumed to be separable. In particular, the space $L^2(E)$, where E is a measurable set, automatically meets this requirement. For Hilbert space that is not separable, you may refer to Problem 2, which can be better understood after grasping the idea presented in Exercise 4.a and 4.b of the Section 4.3.

4.1 Exercises

Exercise 4.8 Let $\eta(t)$ be a fixed positive (measurable) function real-valued on $[a,b]$. Define $\mathcal{H}_\eta = L^2([a,b],\eta)$ to be the space of all functions f on $[a,b]$ such that

$$\int_a^b |f(t)|^2\eta(t)\ \mathrm{d}t < \infty.$$

Define the inner product on $\mathcal{H}_\eta$ by

$$(f,g)_\eta = \int_a^b f(t)\overline{g(t)}\eta(t)\ \mathrm{d}t.$$

Show that $\mathcal{H}_\eta$ is a Hilbert space and that the mapping $U : f \mapsto \eta^{1/2}f$ gives a unitary correspondence between $\mathcal{H}_\eta$ and the usual space $L^2([a,b])$.

Solution Clearly, $\mathcal{H}_\eta$ is a vector space and the inner product is positively definite. A useful observation is that given a function f, $f/\sqrt{\eta}$ is then measurable whenever f is measurable. Also, $f \in \mathcal{H}_\eta$ whenever $f\sqrt{\eta} \in L^2([a,b])$.

To verify that $\mathcal{H}_\eta$ is separable, we use the result of the density of step functions with rational boundaries in $L^2([a,b])$. Suppose $\|f\sqrt{\eta} - g\|_{L^2} < \varepsilon$, and we must have $\|f - g/\sqrt{\eta}\|_{\mathcal{H}_\eta} = \|f\sqrt{\eta} - g\|_{L^2} < \varepsilon$.

Next we check the completeness of $\mathcal{H}_\eta$. Assume $\{f_i\}$ is a Cauchy sequence in $\mathcal{H}_\eta$, then $\{f_i\sqrt{\eta}\}$ is Cauchy in $L^2([a,b])$. Consequently, it converges to a g in $L^2([a,b])$, and f_i definitely converges to $g/\sqrt{\eta}$ in $\mathcal{H}_\eta$.

That U is a unitary mapping is a direct outcome of the two observations aforementioned. ■

Exercise 4.9 Let $\mathcal{H}_1 = L^2([-\pi, \pi])$ be the Hilbert space of functions $F(e^{i\theta})$ on the unit circle with inner product $(F, G) = \frac{1}{2\pi}\int_{-\pi}^{\pi} F(e^{i\theta})\overline{G(e^{i\theta})}\,\mathrm{d}\theta$. Let $\mathcal{H}_2$ be the space $L^2(\mathbb{R})$. Using the mapping

$$x \mapsto e^{2i\arctan x} \qquad \left(\text{which equals } x \mapsto \frac{i-x}{i+x}\right)$$

of $\mathbb{R}$ to the unit circle, show that:

(a) The correspondence $U : F \to f$, with

$$f(x) = \frac{1}{\pi^{1/2}(i+x)} F\left(\frac{i-x}{i+x}\right)$$

gives a unitary mapping of $\mathcal{H}_1$ to $\mathcal{H}_2$.

(b) As a result,

$$\left\{\frac{1}{\pi^{1/2}}\left(\frac{i-x}{i+x}\right)^n \frac{1}{i+x}\right\}_{n=-\infty}^{\infty}$$

is an orthonormal basis of $L^2(\mathbb{R})$.

Hint As for (a), you may recall the change of variable formula (Equation 3.3) and you can construct yourself other types of unitary mapping, such as

$$f(x) = \frac{1}{\sqrt{\pi(1+x^2)}}\, F\left(\frac{i-x}{i+x}\right).$$

Yet the correspondence and the mapping listed in this exercise is most naturally arisen in scientific research.

Once you prove it, by noting that unitary mapping maps an orthonormal basis to an orthonormal basis, what is one such basis of $\mathcal{H}_1$?

Exercise 4.10 Let $\mathcal{S}$ denote a subspace of a Hilbert space $\mathcal{H}$. Prove that $(\mathcal{S}^\perp)^\perp$ is the smallest closed subspace of $\mathcal{H}$ that contains $\mathcal{S}$.

Solution Clearly, every intersection of closed subspaces is still a closed subspace. Hence the "smallest" is well-defined. Let $\overline{\mathcal{S}} \supseteq \mathcal{S}$ be the smallest closed subspace, and we must have $\overline{\mathcal{S}} \subseteq (S^\perp)^\perp$, since $(S^\perp)^\perp$ contains $\mathcal{S}$ and is closed. Conversely, there is $\overline{\mathcal{S}}^\perp = \mathcal{S}^\perp$. It can be reasoned as follows: Suppose $x \in \mathcal{S}^\perp$, then $\overline{\mathcal{S}} \subseteq x^\perp$, which implies that $x \in \overline{\mathcal{S}}^\perp$. Moreover, $\overline{\mathcal{S}}^\perp \subseteq \mathcal{S}^\perp$ naturally holds. Note that

$$\mathcal{S}^\perp \oplus (\mathcal{S}^\perp)^\perp = \mathcal{H} = \overline{\mathcal{S}} \oplus \overline{\mathcal{S}}^\perp = \overline{\mathcal{S}} \oplus \mathcal{S}^\perp \text{ and that } \overline{\mathcal{S}} \subseteq (\mathcal{S}^\perp)^\perp,$$

and we must have $(\mathcal{S}^\perp)^\perp \subseteq \overline{\mathcal{S}}$. ■

Remark A direct corollary is that given $\mathcal{S} \subseteq \mathcal{H}$ is closed, $(\mathcal{S}^\perp)^\perp = \mathcal{S}$.

Exercise 4.10' In this exercise, we investigate the closure of a subspace $\mathcal{S}$ of the Hilbert space $\mathcal{H}$.

(a) Just like what we know of the closure in $\mathbb{R}^d$, we define

$$\mathcal{S}_1 = \{h \in \mathcal{H} : \text{ there exists a sequence } \{h_i\}_{i\in\mathbb{N}} \text{ in } \mathcal{S} \text{ such that } h_i \to h \text{ in } \mathcal{H}\}.$$

Prove that $\overline{\mathcal{S}} = \mathcal{S}_1$. A similar methodology is embedded in the completion procedure of a pre-Hilbert space.

(b) Suppose $\{e_i\}_{i=1}^{\infty}$ is an orthonormal basis of $\mathcal{H}$ and $S \subseteq \mathbb{N}_+$. Show that the smallest closed subspace

containing $\{e_i\}_{i\in S}$ is exactly
$$\left\{\sum_{i\in S} a_i e_i : \sum_{i\in S} |a_i|^2 < \infty\right\}.$$

(c) Given two subspaces $\mathcal{S}, \mathcal{T} \subseteq \mathcal{H}$ with $\mathcal{S} \perp \mathcal{T}$ but are not necessarily closed. Then $\overline{\mathcal{S}} \perp \overline{\mathcal{T}}$.

(d) Suppose $\{\mathcal{S}_i\}_{i=1}^{\infty}$ is a sequence of *closed* subspaces of $\mathcal{H}$ which is pairwise perpendicular, i.e., for $i \neq j$, $\mathcal{S}_i \perp \mathcal{S}_j$. Prove that $\sum_{i=1}^{\infty} \mathcal{S}_i$ is direct. Furthermore, $\bigoplus_{i=1}^{\infty} \mathcal{S}_i$ is a closed subspace of $\mathcal{H}$ as well. In case of countably infinite summation, $\sum_{i=1}^{\infty} \mathcal{S}_i$ is defined as
$$\left\{\sum_{i=1}^{\infty} s_i : \sum_{i=1}^{\infty} \|s_i\|^2 < \infty, \text{ while } s_i \in \mathcal{S}_i \text{ for any } i \in \mathbb{Z}_+\right\}.$$

Hint

(a) Clearly, $\mathcal{S} \subseteq \mathcal{S}_1$. Show that any sequence $\{h_i\}_{i\in\mathbb{N}}$ in $\mathcal{S}_1$ that converges to a point h implies that $h \in \mathcal{S}_1$.

(b) Deduce it directly from Theorem 2.3 in the textbook.

(c) Use (a).

(d) That the sum is direct is by definition. Consider a Cauchy sequence $\{\sum_{i=1}^{\infty} s_{i,j}\}_{j=1}^{\infty}$. Since $\|s_{i,j} - s_{i,j'}\| \leq \|\sum_{i=1}^{\infty} s_{i,j} - s_{i,j'}\|$, we have that there exists $s^{(i)} \in \mathcal{S}_i$ such that $s_{i,j} \to s^{(i)}$ as j goes to infinity. Next we claim that $\sum_{i=1}^{\infty} \|s^{(i)}\|^2 \leq \infty$. Given $\varepsilon > 0$, for M, K sufficiently large, momentarily fix N and we must have
$$\varepsilon > \left\|\sum_{i=1}^{\infty} s_{i,M} - s_{i,K}\right\|^2 \geq \left\|\sum_{i=1}^{N} s_{i,M} - s_{i,K}\right\|^2 = \sum_{i=1}^{N} \|s_{i,M} - s_{i,K}\|^2. \tag{4.1}$$
Letting K goes to infinity, we have $\sum_{i=1}^{N} \|s_{i,M} - s^{(i)}\|^2 \leq \varepsilon$. Denote $\sup_{j\in\mathbb{N}} \|\sum_{i=1}^{\infty} s_{i,j}\|^2 < \infty$ as L and we see that
$$\sum_{i=1}^{N} \|s^{(i)}\|^2 \leq 2\sum_{i=1}^{N} \|s_{i,M} - s^{(i)}\|^2 + \|s_{i,M}\|^2 \leq 2(L + \varepsilon).$$
Choosing ε arbitrarily close to 0 and then picking N sufficiently large, we have proved the claim. Finally, we are to establish the convergence to $\sum_{i=1}^{\infty} s^{(i)}$. In Equation 4.1, letting $K \to \infty$ and then $N \to \infty$, we have
$$\left\|\sum_{i=1}^{\infty} s_{i,M} - s^{(i)}\right\|^2 \leq \varepsilon$$
for M great enough. ■

Exercise 4.11 Let P be the orthogonal projection associated with a closed subspace $\mathcal{S}$ in a Hilbert space $\mathcal{H}$.

(a) Show that $P^2 = P$ and $P^* = P$.

(b) Conversely, if P is any bounded operator satisfying $P^2 = P$ and $P^* = P$, prove that P is the orthogonal projection for some closed subspace of $\mathcal{H}$.

(c) Using P, prove that if $\mathcal{S}$ is a closed subspace of a separable Hilbert space, then $\mathcal{S}$ is also a separable Hilbert space.

Solution

(a) Given a vector v in $\mathcal{H}$, denote $v_1 = P(v)$ and $v_2 = v - v_1$ be the component parallel to $\mathcal{S}$ and vertical to $\mathcal{S}$ respectively. Then $P^2(v) = P(v_1) = v_1$ and

$$(P(x), y) = (x_1, y_1 + y_2) = (x_1, y_1) = (x_1 + x_2, y_1) = (x, P(y)).$$

(b) Let $\mathcal{S} := \operatorname{Ker} P$ and $\mathcal{T} = \operatorname{Im} P$ be two subspaces. Due to the continuity nature of P, $\mathcal{S}$ is closed. By noting that $x = (x - Px) + Px \in \mathcal{S} + \mathcal{T}$, we must have $\mathcal{H} = \mathcal{S} + \mathcal{T}$. Furthermore, $\mathcal{S} \cap \mathcal{T} = \mathbf{0}$. Indeed, let $x \in \mathcal{S} \cap \mathcal{T}$, then $x = Px = 0$. It only remains to show that $\mathcal{T} \subseteq \mathcal{S}^{\perp}$. Actually, given $t \in \mathcal{T}$, there is

$$0 = (Ps, t) = (s, Pt) = (s, t) \quad \text{for any } s \in \mathcal{S}.$$

(c) Suppose $\{e_i\}_{i=1}^{\infty}$ is one separable family of $\mathcal{H}$. Consider the orthogonal projection onto $\mathcal{S}$, which we denote as P, and we show that $\{P(e_i)\}_{i=1}^{\infty}$ is a separable family of $\mathcal{S}$. Given $s \in \mathcal{S}$, let $\|s - \sum_{i=1}^{N} a_i e_i\| < \varepsilon$. We have

$$\left\| s - P\left(\sum_{i=1}^{N} a_i e_i\right) \right\|^2 = \left\| s - \sum_{i=1}^{N} a_i e_i \right\|^2 - \left\| \sum_{i=1}^{N} a_i e_i - P\left(\sum_{i=1}^{N} a_i e_i\right) \right\|^2 < \varepsilon^2.$$

Thereby $\mathcal{S}$ is separable.

■

Exercise 4.13 Suppose P_1 and P_2 are a pair of orthogonal projections onto $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively. Then P_1P_2 is an orthogonal projection if and only if P_1 and P_2 commute. In this case, P_1P_2 projects onto $\mathcal{S}_1 \cap \mathcal{S}_2$.

Solution It is a paradigm of how carfully specified properties will come to save us from getting bogged down with direct yet lengthy discussion of all possible cases. Indeed, if P_1 and P_2 commute, for the sake of showing that P_1P_2 is another orthogonal projections we only have to verify that $P_1P_2 = P_2P_1$ and that $(P_1P_2)^* = P_1P_2$. [Note that P_1P_2 is sure to be bounded.]

For the other direction, we have $P_2P_1 = P_2^*P_1^* = (P_1P_2)^* = P_1P_2$. Also, by noting that $\operatorname{Im} P_1P_2 \subseteq \mathcal{S}_1$ and that $\operatorname{Im} P_2P_1 \subseteq \mathcal{S}_2$, we have $\operatorname{Im} P_1P_2 \subseteq \mathcal{S}_1 \cap \mathcal{S}_2$. But it is clear that $\mathcal{S}_1 \cap \mathcal{S}_2 \subseteq \operatorname{Im} P_1P_2$. So P_1P_2 must be the projection onto $\mathcal{S}_1 \cap \mathcal{S}_2$. ■

Exercise 4.18 Let $\mathcal{H}$ denote a hilbert space, and $\mathcal{L}(\mathcal{H})$ the vector space of all bounded linear operators on $\mathcal{H}$. Indeed, $\mathcal{L}(\mathcal{H})$ forms a ring w.r.t. addition, scalar multiplication and functional composition. However, we will not concern composition for now.

(a) Prove that

$$d(T_1, T_2) = \|T_1 - T_2\|$$

defines a metric on $\mathcal{L}(\mathcal{H})$.

(b) Show that $\mathcal{L}(\mathcal{H})$ is complete in measure d.

Solution We intentionally omit the explanation of (a) and target directly at (b). This form of completeness has quite a disparate appearance compared with that of L^1 and L^2.

Suppose $\{T_i\}_{i\in\mathbb{N}}$ is Cauchy in d. For a given $v \in \mathcal{H}$, note that $\{T_i(v)\}_{i\in\mathbb{N}}$ is Cauchy in $\mathcal{H}$, which inspires us to define a function $T : v \mapsto \lim_{i\to\infty} T_i(v)$. Clearly, T is linear because of the linearity of T_i. Moreover, $T \in \mathcal{L}(\mathcal{H})$. The reason is simple: Select a sub-sequence $\{T_{i_k}\}_{k=1}^{\infty}$ such that $\|T_{i_{k+1}} - T_{i_k}\| \leq 2^{-k}$ for all k. We must have

$$\|T(v)\| = \left\| T_{i_1}(v) + \sum_{k=1}^{\infty} (T_{i_{k+1}} - T_{i_k})(v) \right\| \le \|T_{i_1}(v)\| + \sum_{k=1}^{\infty} \|(T_{i_{k+1}} - T_{i_k})(v)\| \le (B_{i_1} + 1)\|v\|.$$

Using a similar argument, we can prove that $d(T_i, T) \to 0$ as $i \to \infty$. ■

Exercise 4.19 If T is a bounded linear operator on a Hilbert space, prove that

$$\|TT^*\| = \|T^*T\| = \|T\|^2 = \|T^*\|^2.$$

Solution Undoubtedly, $\|TT^*\| = \|T^*T\| \le \|T\|^2 = \|T^*\|^2$ holds. Let M denote $\|T\|$ and we can select $\|f\| = 1$ such that $\|T^*f\| > 1 - \varepsilon$. Hence, $(TT^*f, f) = (T^*f, T^*f) > (1-\varepsilon)^2$. ■

Exercise 4.20 (Weak Convergence in $\mathcal{H}$) Suppose $\mathcal{H}$ is an infinite-dimensional Hilbert space. Show that $\mathcal{H}$ enjoys a notion of **weak convergence**, i.e., for any sequence $\{f_n\}$ in $\mathcal{H}$ with $\|f_n\| = 1$ for all n, there exists $f \in \mathcal{H}$ and a subsequence $\{f_{n_k}\}$ such that for all $g \in \mathcal{H}$, one has

$$\lim_{k\to\infty} (f_{n_k}, g) = (f, g).$$

Solution WLOG, assume that $\mathcal{H}$ is $\ell^2(\mathbb{Z})$ and let e_i denote the one-hot vector with i-th coordinates one. Since $0 \le |(f_n, e_0)| \le 1$ for all n and i, then one divide the square $[-1,1] \times [-1,1]$ into four pieces $S_{1,k}$ $(1 \le k \le 4)$ of equal size. Then one of $\{f_n : (f_n, e_0) \in S_{1,k}\}$ is infinite, and denote that one as the set S_0. Since S_0 is infinite, then sub-divide it into 64 portions depending on which pair of k and l satisfies $(f_n, e_{-1}) \in S_{1,k}$ and $(f_n, e_1) \in S_{1,l}$ and which quarter (f_n, e_0) lies in the previous square. Select one subset among them with infinite elements and denote it as S_1. Continuing this process indefinitely, every time we divide each of the previous intervals into four and add to it a pair of new coordinates, and then select the one subset with infinite cardinality. After that we select an arbitrary element with index n_0 from S_0, and since S_1 has infinitely many elements, we can always select from S_1 an element with index $n_1 > n_0$ and so forth. We shall use the sequence $\{f_{n_k}\}$ derived as the one desired in the conclusion.

We next claim that for all e_i, $a_i := \lim_{k\to\infty}(f_{n_k}, e_i)$ exists. The reason is straightforward: (f_{n_k}, e_i) forms a Cauchy sequence by its construction. Most importantly, $\sum_{i\in\mathbb{Z}} |a_i|^2 \le 1$. Consider the partial sum $\sum_{i=-N}^{N} |a_i|^2$ here. Since $\sum_{i=-N}^{N} |(f_{n_k}, e_i)|^2 \le \|f_{n_k}\|^2 = 1$, by taking k sufficiently large, we have $\sum_{i=-N}^{N} |a_i|^2 \le 1$. Then letting N go to infinity, we have $\sum_{i\in\mathbb{Z}} |a_i|^2 \le 1$.

We may define $f = \sum_{i=-\infty}^{\infty} a_i e_i$, which is of course well-defined due to previous discussion. Now that $g = e_i$ satisfies $\lim_{k\to\infty}(f_{n_k}, g) = (f, g)$, any finite combination of e_i is satisfying too. To prove it for all $g \in \mathcal{H}$, we recall that $\{e_i\}$ is dense in $\mathcal{H}$. Suppose g' is a finite linear combination such that $\|g - g'\| \le \varepsilon$, and we have

$$\begin{aligned}\limsup_{k\to\infty} |(f_{n_k}, g) - (f, g)| &\le \sup |(f_{n_k}, g - g')| + \lim_{k\to\infty} |(f_{n_k}, g') - (f, g')| + |(f, g - g')| \\ &\le 2\|g - g'\| + \|g - g'\| < 3\varepsilon.\end{aligned}$$

Choosing ε arbitrarily small, we have $\lim_{k\to\infty}(f_{n_k}, g) = (f, g)$ for all $g \in \mathcal{H}$. ■

Remark The procedure above is able to provide a Cauchy sequence once there exists a $g \in \mathcal{H}$ such that each of the coordinates can be bounded by that of g. See Exercise 4.24.

Exercise 4.21 (Notions of Convergence in $\mathcal{L}(\mathcal{H})$) In Exercise 4.18 we introduced the notion of convergence in the norm. Next, there is a weaker convergence, which happens to be called **strong convergence**, that requires that $T_n f \to Tf$, as $n \to \infty$, for every vector $f \in \mathcal{H}$. Finally, there is **weak convergence** (See also Exercise 4.20) that requires $(T_n f, g) \to (Tf, g)$ for every pair of vectors $f, g \in \mathcal{H}$.

(a) Show by examples that weak convergence does not imply strong convergence nor does strong convergence imply convergence in the norm.

(b) Show that for any bounded operator T there is a sequence $\{T_n\}$ of bounded operators of finite rank so that $T_n \to T$ strongly as $n \to \infty$.

Solution Let $\{e_i\}_{i=1}^{\infty}$ be a orthonormal basis of $\mathcal{H}$ and $S_k = \langle e_i \rangle_{i=1}^{k}$ the (closed) subspace spanned by the first k bases. Also, denote the null operator as $\mathcal{O}$ and the orthogonal projection onto $S_k^{\perp}$ as P_k. Finally, let

$$P : \sum_{i=1}^{\infty} a_i e_i \mapsto \sum_{i=1}^{\infty} a_i e_{i+1}$$

be a bounded linear transformation on $\mathcal{H}$.

First we claim that $P_n \to \mathcal{O}$ strongly, as $n \to \infty$. In fact, $\|P_n f\| = \sum_{k>n} |(f, e_i)|^2 \to 0$. However, $\|P_n - \mathcal{O}\| = 1$ for all $n \in \mathbb{N}$.

Next we show that $P^n \to \mathcal{O}$ weakly. Indeed,

$$|(P^n f, g)| = |(P^n f, P_n g)| \leq \|P^n f\| \|P_n g\| = \|f\| \|P_n g\| \to 0.$$

Letting $f = e_1$, we see that $P^n f = e_{n+1}$, so P^n cannot converge strongly.

As for (b), picking T_n as $(\mathcal{I} - P_n)T$ satisfies all the requirements, where $\mathcal{I}$ denotes the identity. ■

Exercise 4.22 An operator is an **isometry** if $\|Tf\| = \|f\|$ for all $f \in \mathcal{H}$.

(a) Show that if T is an isometry, then $(Tf, Tg) = (f, g)$ for every $f, g \in \mathcal{H}$. Prove as a result that $T^*T = \mathcal{I}$.

(b) If T is an isometry and T is surjective, then T is unitary and $TT^* = \mathcal{I}$.

(c) If T^*T is unitary, then T is an isometry.

Solution (a) is a simple application of the polarization identity

$$(f, g) = \frac{1}{4}\left[\|f + g\|^2 - \|f - g\|^2 + i\|f + ig\|^2 - i\|f - ig\|^2\right].$$

Also note that T is automatically injective.

For (b), of course T is bijective and it follows that T is unitary. Now that $T^* = T^{-1}$, we have $TT^* = \mathcal{I}$ as desired.

Frankly speaking, I discovered no simple solution to (c). Here I provide my solution to this problem which is a bit involved, but fortunately we will get inspiring intermediate result along the way.

Lemma 4.1 (Reflection)

Suppose A is a bounded linear operator satisfying $A^2 = \mathcal{I}$, we have $\mathcal{H} = V_1 \oplus V_{-1}$, where $V_\lambda := \mathrm{Ker}(A - \lambda\mathcal{I})$ is a closed subspace of $\mathcal{H}$. Moreover, if A is unitary, then $A = P_1 - P_{-1}$, where P_λ denotes the orthogonal projections onto V_λ. ♡

Proof For any $f \in \mathcal{H}$, decompose f as $\frac{1}{2}[(A - \mathcal{I})f - (A + \mathcal{I})f]$, and we have $\mathcal{H} = V_1 + V_2$. Also, the sum is direct, since if $f \in V_1 \cap V_{-1}$ there is $f = Af = -Af = -f$. If A is unitary, given $f \in V_1$ and $g \in V_2$, we have $(f,g) = (Af, Ag) = (f, -g) = -(f,g)$, so $V_2 \subseteq V_1^{\perp}$ and hence $V_2 = V_1^{\perp}$. It is clear at this point that $A = P_1 - P_{-1}$. □

Back to the original problem.

Since $(T^*T)^2 = \mathcal{I}$ and T^*T is unitary, we have $T^*T = P_{\mathcal{S}} - P_{\mathcal{S}^{\perp}}$ by the lemma. Next we claim that $\mathcal{S} = \mathcal{H}$. Otherwise, suppose $0 \neq g \in \mathcal{S}^{\perp}$. We have

$$\|Tg\|^2 = (T^*Tg, g) = -\|g\|^2 < 0.$$

This contradiction shows $T^*T = \mathcal{I}$. It then follows that $(f,g) = (T^*Tf, g) = (Tf, Tg)$ for every $f, g \in \mathcal{H}$. This completes the proof. ■

Exercise 4.23 Suppose $\{T_k\}$ is a collection of bounded operators on a Hilbert space $\mathcal{H}$, with $\|T_k\| \leq 1$ for all k. Suppose also that

$$T_k^* T_j = T_k T_j^* = 0 \quad \text{for all } k \neq j.$$

Let $S_N = \sum_{-N}^{N} T_k$.

Show that S_N converges strongly as $N \to \infty$. Moreover, the limiting linear operator has a norm ≤ 1. A direct generalization is given in Problem 4.8 below.

Hint The crux lies in the case where there are only two operators. Can you prove it? Once proved, you can generalize it to a coutably infinite collection using conclusions in Exercise 4.10'.

Solution We shall denote the range of T_i^* as $\operatorname{Im} T_i^*$ and its closure as $\mathcal{S}_i^*$. Several observations are in order.

1. $\operatorname{Im} T_i^* \perp \operatorname{Im} T_j^*$ whenever $i \neq j$. By (c) in Exercise 4.10', we have $\mathcal{S}_i^* \perp \mathcal{S}_j^*$.
2. Clearly, $\operatorname{Ker} T_i \supseteq \operatorname{Im} T_j^*$ whenever $i \neq j$. Since $\operatorname{Ker} T_i$ is closed, there is $\operatorname{Ker} T_i \supseteq \mathcal{S}_j^*$.
3. Let $\mathcal{T} = \left(\bigoplus_{i=-\infty}^{\infty} \mathcal{S}_i^*\right)^{\perp}$, which is well-defined according to (d) in Exercise 4.10'. Given any $x \in \mathcal{H}$, we may conveniently write
$$x = t + \sum_{i=-\infty}^{\infty} s_i, \qquad \text{where } t \in \mathcal{T} \text{ and } s_i \in \mathcal{S}_i^* \text{ for all } i,$$
such that $\sum_{i=-\infty}^{\infty} \|s_i\|^2 < \infty$.
4. $\mathcal{T} \subseteq \operatorname{Ker} T_i$ for any i. Note that $(T_i t, u) = (t, T_i^* u) = 0$ for any $u \in \mathcal{H}$.
5. Due to the continuity of T_i, we must have $T_j x = T_j t + \sum_{i=-\infty}^{\infty} T_j s_i = T_j s_j$.

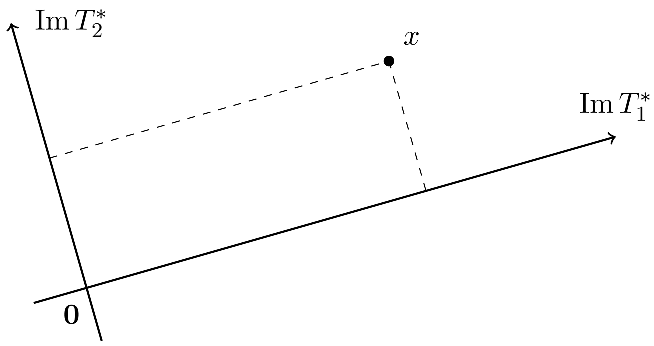

Figure 4.1: Orthogonal decomposition of the vector $x \in \mathcal{H}$

Equiped with these observations, we have

$$\begin{aligned}
\|x\|^2 &= \|t\|^2 + \sum_{i=-\infty}^{\infty} \|s_i\|^2 && \triangleright \text{ Orthogonal decomposition} \\
&\geq \sum_{i=-\infty}^{\infty} \|T_i s_i\|^2 && \triangleright \|T_i\| \leq 1 \text{ for all } i \\
&= \sum_{i=-\infty}^{\infty} \|T_i x\|^2 && \triangleright \text{ Observation 5} \\
&= \left\| \left(\sum_{i=-\infty}^{\infty} T_i \right) x \right\|^2 = \|Tx\|^2 && \triangleright \left\{ \textstyle\sum_{i=-N}^{N} T_i\, x \right\} \text{ is Cauchy}
\end{aligned}$$

It can be immediately verified that T is a bounded linear operator. Moreover, $\|T\| \leq 1$, as is implied by the estimation above. ■

Exercise 4.24 can be handled by the same construction procedure as in Exercise 4.20.

Exercise 4.25 (Truncating Argument) Suppose T is a bounded operator that is diagonal with respect to a basis $\{\varphi_k\}$, with $T\varphi_k = \lambda_k \varphi_k$. Then T is compact iff $\lambda_k \to 0$.

Solution For the necessity side, take the sequence $\{\varphi_k\}$. For the sufficiency side, let P_n denote the orthogonal projections onto $\langle \varphi_1, \ldots, \varphi_n \rangle$, which surely is compact. Note that $\|P_n T - P\| = \sup_{k>n} |\lambda_k| \to 0$, as $n \to \infty$. Since the compact operator is closed under convergence in the norm, we have that T is compact. ■

Exercise 4.26 (Integral Operators in General) Suppose w is a measurable function on $\mathbb{R}^d$ with $0 < w(x) < \infty$ for a.e. x, and K is a measurable function on $\mathbb{R}^{2d}$ that satisfies:

(i) $\displaystyle\int_{\mathbb{R}^d} |K(x,y)| w(y)\ \mathrm{d}y \leq A w(x)$ for a.e. $x \in \mathbb{R}^d$, and

(ii) $\displaystyle\int_{\mathbb{R}^d} |K(x,y)| w(x)\ \mathrm{d}x \leq A w(y)$ for a.e. $y \in \mathbb{R}^d$.

Prove that the integral operator defined by

$$Tf(x) = \int_{\mathbb{R}^d} K(x,y) f(y)\ \mathrm{d}y, \quad x \in \mathbb{R}^d$$

is bounded on $L^2(\mathbb{R}^d)$ with $\|T\| \leq A$.

Solution Here w is used to balance the mass of K. By the Cauchy-Schwarz inequality, we must have

$$\begin{aligned}
|Tf(x)| &\leq \int_{\mathbb{R}^d} |K(x,y)||f(y)|\ \mathrm{d}y \\
&\leq \left(\int_{\mathbb{R}^d} |K(x,y)| w(y)\ \mathrm{d}y \right)^{1/2} \left(\int_{\mathbb{R}^d} |K(x,y)||f(y)|^2 w(y)^{-1}\ \mathrm{d}y \right)^{1/2} \\
&\leq A^{1/2} w(x)^{1/2} \left(\int_{\mathbb{R}^d} |K(x,y)||f(y)|^2 w(y)^{-1}\ \mathrm{d}y \right)^{1/2}.
\end{aligned}$$

Thus, the Fubini's theorem yields

$$\begin{aligned}
\|Tf\|^2 &= \int_{\mathbb{R}^d} |Tf(x)|^2\ \mathrm{d}x \\
&\leq A \int_{\mathbb{R}^d} w(x) \left(\int_{\mathbb{R}^d} |K(x,y)||f(y)|^2 w(y)^{-1}\ \mathrm{d}y \right)\ \mathrm{d}x
\end{aligned}$$

$$
\begin{aligned}
&= A \int_{\mathbb{R}^d} w(y)^{-1} \left(\int_{\mathbb{R}^d} |K(x,y)| w(x) \, \mathrm{d}x \right) |f(y)|^2 \, \mathrm{d}y \\
&\le A^2 \|f\|^2,
\end{aligned}
$$

which not only shows that T is an operator on $L^2(\mathbb{R}^d)$, but $\|T\| \le A$ as well. ■

Exercise 4.27 Prove that the operator

$$Tf(x) = \frac{1}{\pi} \int_0^\infty \frac{f(y)}{x+y} \, \mathrm{d}y$$

is bounded on $L^2(0,\infty)$ with norm $\|T\| \le 1$.

Solution I found w by a mere chance of serendipity. However, I believe that there must be some insights. If you do have any idea, please drop me a line.

Let $w(x) = x^{-1/2}$. We may verify the condition stipulated in Exercise 4.26 as follows:

$$
\begin{aligned}
\int_0^\infty \frac{1}{\sqrt{x}(x+y)} \, \mathrm{d}x &= \frac{2}{\sqrt{y}} \int_0^\infty \frac{\mathrm{d}t}{(t^2+1)} && \triangleright \text{ Substitute } x = yt^2 \\
&= \frac{\pi}{\sqrt{y}} = \pi \cdot w(y).
\end{aligned}
$$

Therefore $\|T\| \le 1$. ■

Exercise 4.28 Suppose $\mathcal{H} = L^2(B)$, where B is the unit ball in $\mathbb{R}^d$. Let $K(x,y)$ be a measurable function on $B \times B$ that satisfies $|K(x,y)| \le A|x-y|^{-d+\alpha}$ for some $\alpha > 0$, whenever $x, y \in B$. Define

$$Tf(x) = \int_B K(x,y) f(y) \, \mathrm{d}y.$$

(a) Prove that T is a bounded operator.
(b) Prove that T is compact.
(c) Note that T is *guaranteed* to be a Hilbert-Schmidt operator iff $\alpha > d/2$.

Solution

(a) It closely follows the technique presented in Exercise 4.26 above. Since

$$\int_B |K(x,y)| \, \mathrm{d}y \le A \int_{2B} |y|^{-d+\alpha} \, \mathrm{d}y < \infty,$$

we may choose $w \equiv 1$.

(b) Let

$$K_n(x,y) = \begin{cases} |K(x,y)| & \text{if } |x-y| > \frac{1}{n} \\ 0 & \text{otherwise.} \end{cases}$$

and T_n be the integral operator associated with this truncated kernels. Then T_n is Hilbert-Schmidt. Moreover, invoking the technique one more time with $w \equiv 1$, we see that

$$\|T_n - T\| \le A \int_{\frac{2}{n}B} |y|^{-d+\alpha} \, \mathrm{d}y \to 0.$$

Due to the fact that the convergence in the norm perseves compactness, we have therefore T is compact.

(c) If $\alpha > d/2$, then clearly $K \in L^2(B)$. For the reverse order, let $K(x,y) = A|x-y|^{-d+\alpha}$. We need the following lemma to confirm the uniqueness of the kernel. Though stated in a conservative way, it should be gratifying for now. Note that $L^2(B) \subseteq L^1(B)$. Up to this moment, we can say nothing more than $K(x,y) \in L^1(B)$ of K.

Lemma 4.2 (Uniqueness of the Kernel)

Suppose that $K(x,y) \in L^1(B)$ is a non-negative function defined on $B \times B$ such that the integral operator T associated with K maps $L^2(B)$ to $L^2(B)$. If T is at the same time a Hilbert-Schmidt operator whose kernel is $M(x,y)$, then $K(x,y) \in L^2(B \times B)$ and $K(x,y) = M(x,y)$ almost everywhere.

♡

Proof Let $\{e_i\}$ be a denumerable orthonormal basis of $L^2(B)$ and $K_x(y)$ the slice induced by fixing x. It suffices to show that if T equals the null operator, then $K = 0$ a.e..
Indeed, given any i, Te_i is non-zero (either $\neq 0$ or ill-defined) on a null set of B, which we denote as B_i. Clearly, $m(\bigcup_i B_i) = 0$. Hence, for every $x \in \bigcap_i \overline{B_i}$ we have $\int_{B_x} K_x(y) f(y) \, \mathrm{d}y = 0$ for all $f \in \{e_i\}$. We may choose $\{e_i\}$ to be non-negative functions and constrain our attention to f being non-negative as well. By Fatou's lemma, decompsing f as possibly infinite sum of $\{e_i\}$, we must have

$$\int_{B_x} K_x(y) f(y) \, \mathrm{d}y \leq \liminf_{N \to \infty} \int_{B_x} K_x(y) \sum_{i=1}^{N} (f, e_i) e_i(y) \, \mathrm{d}y = 0.$$

Choosing appropriate f on the following sets: $\{y : K_x(y) > L\}$ when $L > 0$, we may conclude that K_x vanishes a.e.. Using the Fubini's theorem, we conclude that K is a.e. zero and is thus in $L^2(B \times B)$. □

Back to the original problem. By invoking the lemma above, K must be in $L^2(B \times B)$, which directly implies that $\alpha > d/2$.

■

Exercise 4.29 (Fredholm, Continued in Problem 4.6) Let T be a compact operator on a Hilbert space $\mathcal{H}$, and assume $\lambda \neq 0$. [λ does not need to be an eigenvalue.]

(a) Show that the range of $\lambda\mathcal{I} - T$ is closed.
(b) Show by example that this may fail when $\lambda = 0$.
(c) Show that $\mathrm{Im}(\lambda\mathcal{I} - T)$ is all of $\mathcal{H}$ iff $\mathrm{Ker}(\overline{\lambda}\mathcal{I} - T^*)$ is trivial.

Solution

(a) It will be much more tractable if you do (b) first. [At least for me :)] So if you have not tackled this problem, please pause and switch to (b) first, and then continue your reading.
By definition, we only have to show that given a converging sequence $g_k \to g$, where $g_k = (\lambda\mathcal{I} - T) f_k$, we can always find an $f \in \mathcal{H}$ such that $g = (\lambda\mathcal{I} - T) f$. To avoid notational clutter, let $\mathcal{S}$ be $\mathrm{Ker}(\lambda\mathcal{I} - T)$. Clearly, $\mathcal{S}$ is closed according to the continuity of T. WLOG, we assume $f_k \in \mathcal{S}^\perp$.
What we aim to achieve is to demonstrate the boundedness when $\{f_k\}$. Once demonstrated, we may well assume that $Tf_k \to f$. Hence $f_k = \lambda^{-1}(g_k + T_k f_k) \to \lambda^{-1}(g + f) =: f^*$. As a result,

$$(\lambda\mathcal{I} - T) f^* = (f + g) - f = g.$$

Let us delve into the consequences if $\{f_k\}$ is unbounded. By normalizing f_k, due to the bounded nature of g_k, we get a sequence f_k with $\|f_k\| = 1$ and $(\lambda\mathcal{I} - T) f_k \to 0$. If there is a convergent subsequence of $\{f_k\}$, say $f_{n_k} \to \overline{f}$, noting that $\mathcal{S}^\perp$ is closed, we have $\overline{f} \in \mathcal{S}^\perp$. However, $(\lambda\mathcal{I} - T)\overline{f} = 0$, contradicting the definition!

Now we write $Tf_k = \lambda f_k - \delta_k$, where
$$\|\delta_k\| = \|(\lambda\mathcal{I} - T)f_k\| \to 0.$$
According to compactness of T, a subsequence $\{f_{n_k}\}$ exists which makes Tf_{n_k} converges. Then so does $\{f_{n_k} = \lambda^{-1}(Tf_{n_k} + \delta_{n_k})\}$, which is impossible, however.

This completes the proof.

(b) Let $\{e_k\}_{k=1}^{\infty}$ be an orthonormal basis of $\mathcal{H}$. Define
$$T(e_k) = \frac{1}{k}e_k{}^1 \quad \text{for all } k \in \mathbb{Z}.$$
The compactness of T can be verified in a manner reminiscent of Exercise 4.33 below, i.e., by a truncating argument. When it comes to the closeness of $\operatorname{Im} T$, by noting that
$$\operatorname{Im} T = \left\{ \sum_{k=1}^{\infty} a_k e_k : \sum_{k=1}^{\infty} k^2 |a_k|^2 < \infty \right\},$$
we can see that $\sum_{k=1}^{N} \frac{1}{k} e_k \in \operatorname{Im} T$, but the limiting element $\sum_{k=1}^{\infty} \frac{1}{k} e_k$, which of course is well-defined, is not in $\operatorname{Im} T$.

(c) By (a), if we let $\mathcal{S} = \operatorname{Im}(\lambda\mathcal{I} - T)$, then $\mathcal{S}$ is closed. We are to show that $\mathcal{S}^{\perp} = \operatorname{Ker}(\overline{\lambda}\mathcal{I} - T^*)$.

For any $x \in \mathcal{S}^{\perp}$, we know that
$$0 = ((\lambda\mathcal{I} - T)(\overline{\lambda}\mathcal{I} - T^*)x, x) = \|(\overline{\lambda}\mathcal{I} - T^*)x\|^2.$$
Therefore, one side holds: $\mathcal{S}^{\perp} \subseteq \operatorname{Ker}(\overline{\lambda}\mathcal{I} - T^*)$. On the other hand, for any $x \in \operatorname{Ker}(\overline{\lambda}\mathcal{I} - T^*)$, letting f run through $\mathcal{H}$, we have
$$0 = (f, (\overline{\lambda}\mathcal{I} - T^*)x) = ((\lambda\mathcal{I} - T)f, x),$$
which means that $x \in \mathcal{S}^{\perp}$.

This leads to the desired result. Note that $\mathcal{S}^{\perp} = \operatorname{Ker}(\overline{\lambda}\mathcal{I} - T^*)$ may still hold even if $\lambda = 0$.

■

Exercise 4.30 Let $\mathcal{H} = L^2([-\pi, \pi])$ with $[-\pi, \pi]$ identified as the unit circle. Fix a *bounded* sequence $\{\lambda_n\}_{n=-\infty}^{\infty}$ of complex numbers, and define an operator Tf by
$$Tf(x) \sim \sum_{n=-\infty}^{\infty} \lambda_n a_n e^{inx} \quad \text{whenever} \quad f(x) \sim \sum_{n=-\infty}^{\infty} a_n e^{inx}.$$
Such an operator is called a **Fourier multiplier operator**, and the sequence $\{\lambda_n\}$ is called the **multiplier sequence**.

(a) Show that $\|T\| = \sup_n |\lambda_n|$.

(b) Verify that T commutes with translations, that is, if we define $\tau_h(x) = f(x - h)$ then
$$T \circ \tau_h = \tau_h \circ T, \quad \text{for every } h \in \mathbb{R}.$$

(c) Conversely, prove that if T is bounded operator on $\mathcal{H}$ that commutes with translations, then T is a Fourier multiplier operator.

Solution As long as $\{e_n\}$ is an orthonormal basis on which T is diagonalized, T is bounded whenever $\sup_n \|Te_n\| < \infty$. Moreover, $\|T\| = \sup_n \|Te_n\|$.

[1]Note that a bounded linear operator can be uniquely defined by its effect on an orthonormal basis of $\mathcal{H}$.

For (b), we can directly check the identity as follows:

$$\begin{aligned}(T\circ\tau_h)f &= \sum_{n=-\infty}^{\infty}\lambda_n(\tau_h(f),e^{inx})e^{inx}\\ &= \sum_{n=-\infty}^{\infty}\lambda_n(f,e^{in(x+h)})e^{inx}\\ &= \sum_{n=-\infty}^{\infty}\lambda_n(f,e^{inx})e^{in(x-h)}\\ &= \tau_h(Tf)=(\tau_h\circ T)f.\end{aligned}$$

To verify (c), we test the effect of T on e^{inx}. We have

$$\tau_h(T(e^{inx}))=(\tau_h\circ T)e^{inx}=(T\circ\tau_h)e^{inx}=T(e^{in(x-h)})=e^{-inh}T(e^{inx}).$$

Hence $T(e^{inx})$ must be an eigenvector of τ_h for all $h\in\mathbb{R}$ with the eigenvalue e^{-inh}. Let $T(e^{inx})=\sum_{k=-\infty}^{\infty}a_ke^{ikx}$, and then

$$\sum_{k=-\infty}^{\infty}a_ke^{-ikh}e^{ikx}=\sum_{k=-\infty}^{\infty}a_ke^{-inh}e^{ikx}.$$

As a result, $a_k=0$ whenever $k\neq n$. This is exactly what we need to prove. ■

Exercise 4.31 Consider a version of the sawtooth function defined on $[-\pi,\pi)$ by

$$K(x)=i(\operatorname{sgn}(x)\pi-x)\sim\sum_{n\neq 0}\frac{e^{inx}}{n},$$

where the symbol $\operatorname{sgn}(x)$ denotes the sign function. Extend K to $\mathbb{R}$ with period 2π. Suppose $f\in L^1([-\pi,\pi])$ is extended to $\mathbb{R}$ with period 2π as well, and define

$$\begin{aligned}Tf(x)&=\frac{1}{2\pi}\int_{-\pi}^{\pi}K(x-y)f(y)\,\mathrm{d}y\\ &=\frac{1}{2\pi}\int_{-\pi}^{\pi}K(y)f(x-y)\,\mathrm{d}y.\end{aligned}$$

(a) Show that $F(x)=Tf(x)$ is absolutely continuous, and if $\int_{-\pi}^{\pi}f(y)\,\mathrm{d}y=0$, then $F'(x)=if(x)$ a.e. x.

(b) Show that the mapping $f\mapsto Tf$ is compact and symmetric on $L^2([-\pi,\pi])$.

(c) Prove that $\varphi(x)\in L^2([-\pi,\pi])$ is an eigenfunction for T if and only if $\varphi(x)$ is (up to a constant multiple) equal to e^{inx} for some integer $n\neq 0$ with eigenvalue $1/n$, or $\varphi(x)=1$ with eigenvalue 0.

(d) Show as a result that $\{e^{inx}\}_{n\in\mathbb{Z}}$ is an orthonormal basis of $L^2([-\pi,\pi])$.

Solution Indeed, let $S(x,y)=K(x-y)$. Then S is in $L^2([-\pi,\pi]^2)$ and $S(x,y)=\overline{S(y,x)}$, so T is compact and symmetric. Moreover, we may calculate T explicitly:

$$\begin{aligned}(-2\pi i)Tf(x)&=\int_{-\pi}^{\pi}K(y)f(x-y)\,\mathrm{d}y\\ &=\int_0^{\pi}(\pi-y)f(x-y)\,\mathrm{d}y+\int_{-\pi}^{0}(-\pi-y)f(x-y)\,\mathrm{d}y\\ &=\int_{x-\pi}^{x}(\pi-x+t)f(t)\,\mathrm{d}t+\int_{x}^{x+\pi}(-\pi-x+t)f(t)\,\mathrm{d}t \qquad \triangleright\text{ Substitute } t=x-y\end{aligned}$$

$$= \pi \left(\int_{x-\pi}^{x} f(t) \, \mathrm{d}t - \int_{x}^{x+\pi} f(t) \, \mathrm{d}t \right) - x \int_{-\pi}^{\pi} f(t) \, \mathrm{d}t + \int_{x-\pi}^{x+\pi} t f(t) \, \mathrm{d}t.$$

It should be apparent now that T is absolutely continuous. Moreover, the derivative can be computed by

$$(-2\pi i)(Tf)'(x) = 2\pi f(x) - \int_{-\pi}^{\pi} f(t) \, \mathrm{d}t \qquad \triangleright f \text{ is periodic}$$

In order to find eigenvectors of T, we start by noting that 0 is an eigenvalue and its only eigenvectors are constant mappings. Now suppose $Tf = \lambda f$, where $\lambda \neq 0$. Then f is also AC. Taking differentiation, we have

$$-2\pi\lambda i f' = 2\pi f - \int_{-\pi}^{\pi} f(t) \, \mathrm{d}t,$$

which shows f' is AC. Again, letting $g = f'$,

$$-\lambda i g' = g.$$

Construct $h(x) = e^{ix/\lambda} g(x)$, which is AC as well. We know $h' = 0$ a.e., so $g = Ce^{-ix/\lambda}$, where C is a constant. This in turn implies that $f(x) = C_1 e^{ix/\lambda} + C_2$ for some constants C_1, C_2. Substituting f in $Tf = \lambda f$, we can see that λ is the reciprocal of some integers, and the corresponding eigenvectors can be chosen as $e^{ix/\lambda}$. ■

Exercise 4.32 & 4.33 (Necessity of Requirements in the Spectral Theorem)

(4.32) That the operator T is compact is indispensable in the assumption of the spectral theorem. Otherwise, T may have no eigenvectors at all, as is suggested by the following example: Take $T : L^2([0,1]) \to L^2([0,1])$ as

$$Tf(t) = tf(t).$$

It is bounded and symmetric, but not compact. However, T has no eigenvectors.

(4.33) Let $\{e_k\}_{k=1}^{\infty}$ be an orthonormal basis of a Hilbert space $\mathcal{H}$. Verify that the operator defined by

$$T(e_k) = \frac{1}{k} e_{k+1}$$

is compact, but has no eigenvectors.

Solution Obviously, both T defined above are linear operators.

(4.32) Clearly, $T^* = T$. Moreover, $\|T\| = 1$. Next, we show that T is not compact by taking the orthonormal family $\{e^{2\pi inx}\}_{n=0}^{\infty}$ into account. For any $m \neq n$, we have

$$\begin{aligned} \|Te^{2\pi inx} - Te^{2\pi imx}\|^2 &= \int_0^1 x^2 \left| e^{2\pi inx} - e^{2\pi imx} \right|^2 \, \mathrm{d}x \\ &= \int_0^1 x^2 \left(2 - e^{2\pi i(m-n)x} - e^{2\pi i(n-m)x} \right) \, \mathrm{d}x \\ &= \frac{2}{3} - \frac{1}{\pi^2 (m-n)^2}. \qquad \triangleright \text{ Since } \int_0^1 x^2 e^{2\pi ikx} \, \mathrm{d}x = \frac{1}{2\pi ki} + \frac{1}{2\pi^2 k^2} \end{aligned}$$

As a result, no Cauchy subsequence can be found from $\{Te^{2\pi inx}\}_{k=1}^{\infty}$.

Finally, T has no eigenvectors. Suppose $Tf = \lambda f$ for some λ, and then $tf(t) = \lambda f(t)$ for a.e. $t \in [0,1]$. Thus f vanishes a.e..

(4.33) First we show that T has no eigenvectors. Otherwise, let us assume f be the eigenvectors and $i^* = \min\{i \in \mathbb{Z}_+ : (f, e_i) \neq 0\}$. We must have

$$\lambda(f, e_{i^*}) = (Tf, e_{i^*}) = 0.$$

So $\lambda = 0$. However, $Tf = 0$ is by no means possible, since $(Tf, e_{i^*+1}) = \frac{1}{i^*}(f, e_{i^*}) \neq 0$.

Lastly, we briefly discuss the compactness of T. Define

$$T_n(e_i) = \begin{cases} T(e_i) & \text{if } i \leq n, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, T_n is of finite rank and therefore compact. Besides, $\|T_n - T\| = \frac{1}{k+1} \to 0$.

■

Exercise 4.34 Let K be a Hilbert-Schmidt kernel which is real and symmetric. Then, as we saw, the operator T whose kernel is K is compact and symmetric. Let $\{\varphi_k(x)\}$ be the eigenvectors (with eigenvalues λ_k) that digonalize T. Then:

(a) $\sum_k |\lambda_k|^2 < \infty$. $K(x,y) \sim \sum_k \lambda_k \varphi_k(x)\varphi_k(y)$ is the expansion of K in the bases $\{\varphi_k(x)\varphi_y(y)\}$.

(b) Suppose T is a compact operator which is symmetric. Then T is of Hilbert-Schmidt type iff $\sum_k |\lambda_k|^2 < \infty$, where $\{\lambda_k\}$ are the eigenvalues of T counted according to their multiplicities.

Solution

(a) Indeed, we can show that $\sum_k |\lambda_k|^2 = \|K\|^2$.

According to the spectral theorem, we may assume that $\{\varphi_k\}$ is the orthonormal basis. Moreover, since K is real, we can further suppose that φ_k is real. Indeed, decompose $\varphi_k(x) = \varphi_k^{(1)}(x) + i\varphi_k^{(2)}(x)$, and we see that $\varphi_k^{(1)}(x)$ and $\varphi_k^{(2)}(x)$ are both eigenvectors.

From Exercise 4.7 we know that $\{\varphi_k(x)\varphi_j(y)\}$ forms a real-valued orthonormal basis of the domain of K. The (k,j)-th coefficient in the decomposition of K on this basis is

$$\begin{aligned}
(K(x,y), \varphi_k(x)\varphi_j(y)) &= \int_{\mathbb{R}^{2d}} K(x,y)\varphi_k(x)\varphi_j(y)\ \mathrm{d}x\ \mathrm{d}y \\
&= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} K(x,y)\varphi_j(y)\ \mathrm{d}y \right) \varphi_k(x)\ \mathrm{d}x && \triangleright \text{ Fubini's theorem} \\
&= \lambda_j \int_{\mathbb{R}^d} \varphi_j(x)\varphi_k(x)\ \mathrm{d}x \\
& && \triangleright \int_{\mathbb{R}^d} K(x,y)\varphi_j(y)\ \mathrm{d}y = \lambda_j\varphi_j(x) \text{ for a.e. } x \in \mathbb{R}^d \\
&= \lambda_j \delta_{jk}.
\end{aligned}$$

Consequently, $K(x,y) = \sum_k \lambda_k \varphi_k(x)\varphi_k(y)$, where $\sum_k |\lambda_k|^2 = \|K\|^2 < \infty$.

(b) **Hint** We only have to elobrate on the sufficiency, which to some extent is a reverse engineering stuff. The key observation is that guaranteed by the orthonormal nature, we can progressively add new component to our kernel while doing no harm to any other equations beyond our primary concern or satisfied already. The last thing is to show the convergence of the procedure, where we may adopt the notion of convergence in the norm.

Formal Proof

Suppose $\{\varphi_k\}$ is the collection of all eigenvectors with the corresponding eigenvalues $\{\lambda_k\}$. Define the n-th kernel to be

$$K_n(x,y) = \lambda_n \varphi_n(x)\overline{\varphi_n(y)}.$$

Then $(K_n, K_m) = 0$ whenever $n \neq m$. It can also be directly verified that the operator T_n associated with K_n satisfies that

$$T_n(\varphi_n) = \lambda\varphi_n, \text{ and } T_n(\varphi_k) = 0 \text{ for all } k \neq n,$$

which directly implies that $\operatorname{Im} T_n$ is pairwise perpendicular. Moreover,

$$\|T_n\|^2 \le \|K_n\|^2 = \int_{\mathbb{R}^{2d}} K_n(x,y)\overline{K_n(x,y)}\, \mathrm{d}x\, \mathrm{d}y = |\lambda|^2.$$

We have

$$\begin{aligned} \|\sum_{k=n}^{m} T_k\|^2 &\le \sum_{k=n}^{m} \|T_k\|^2 && \triangleright \operatorname{Im} T_i \perp \operatorname{Im} T_j \text{ whenever } i \neq j \\ &\le \sum_{k=n}^{m} |\lambda_k|^2. \end{aligned}$$

According to Exercise 4.18, $\sum_{k=1}^{N} T_k \to T$ in the norm. Besides,

$$\|\sum_{k=n}^{m} K_k\|^2 = \sum_{k=n}^{m} \|K_k\|^2 \le \sum_{k=n}^{m} |\lambda_k|^2,$$

which shows that $\sum_{n=1}^{N} K_n \to K \in L^2(\mathbb{R}^{2d})$. For any $f \in L^2(\mathbb{R}^d)$, now we must have

$$\begin{aligned} Tf = \lim_{N\to\infty} \sum_{n=1}^{N} T_n f &= \lim_{N\to\infty} \int_{\mathbb{R}^{2d}} \sum_{n=1}^{N} K_n(x,y) f(y)\, \mathrm{d}y \\ &= \int_{\mathbb{R}^{2d}} K(x,y) f(y)\, \mathrm{d}y. && \triangleright \text{Dominated convergence theorem} \end{aligned}$$

■

Exercise 4.35 Let $\mathcal{H}$ be a Hilbert space. Prove the following variants of the spectral theorem.

(a) If T_1 and T_2 are two linear and compact operators on $\mathcal{H}$ that commute, show that they can be diagonalized simultaneously. In other words, there exists an orthonormal basis for $\mathcal{H}$ which consists of eigenvectors for both T_1 and T_2.

(b) A linear operator on $\mathcal{H}$ is **normal** if $TT^* = T^*T$. Prove that if T is normal and compact, then T can be diagonalized.

(c) If U is unitary, and $U = \lambda\mathcal{I} - T$, where T is compact, then U can be diagonalized.

Solution

(a) Let V_λ denote $\operatorname{Ker}(T_1 - \lambda\mathcal{I})$, the null-space of $T_1 - \lambda\mathcal{I}$. For any $x \in V_\lambda$, we must have $T_1T_2x = T_2T_1x = \lambda T_2 x$, which shows that V_λ is an invariant subspace of T_2. Restricted on V_λ, $T_2\big|_{V_\lambda}$ is still symmetric and compact, so it can be diagonalized according to the spectral theorem. In particular, we have at least one common eigenvectors for T_1 and T_2.

The rest is almost the same as the main procedure in the proof of the spectral theorem, except that eigenvectors we now consider have to be common.

(b) Inspired by (a), let us write

$$T = \frac{T + T^*}{2} + i\,\frac{T - T^*}{2i} =: T_1 + i\,T_2.$$

Clearly, both T_1 and T_2 is symmetric and compact. Moreover, it can be verified that T_1 and T_2 commute because of $TT^* = T^*T$, so they can simultaneously be diagonalized. Consequently, T is diagonalized with respect to this set of common eigenvectors in turn.

(c) It suffices to show that T can be diagonalized. In fact, $UU^* = \mathcal{I} = U^*U$ when U is unitary, which yields that

$$UU^* = (\lambda\mathcal{I} - T)(\overline{\lambda}\mathcal{I} - T^*) = |\lambda|^2\mathcal{I} - \lambda T^* - \overline{\lambda}T + TT^*,$$

$$U^*U = (\overline{\lambda}\mathcal{I} - T^*)(\lambda\mathcal{I} - T) = |\lambda|^2\mathcal{I} - \lambda T^* - \overline{\lambda}T + T^*T.$$

As a result, $TT^* = T^*T$. Since the conditions in (b) are all satisfied, we deduce that T is diagonalizable. ■

4.2 Problems

Problem 4.2 (Almost Periodic Functions; Non-separable Hilbert Space)

It might be helpful to take a detour by looking at Exercise 4.a and 4.b first. Then proceed to read here. Verify *any* statement in the following narration if you feel the necessity of doing so.

We consider the collection of exponentials $\{e^{i\lambda x}\}_{\lambda\in\mathbb{R}}$. Let $\mathcal{H}_0$ denote the space of finite linear combination of these exponentials. For $f, g \in \mathcal{H}_0$, we define the inner product as

$$(f, g) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(x)\overline{g(x)}\, \mathrm{d}x.$$

With this inner product $\mathcal{H}_0$ is a pre-Hilbert space. Notice that $\|f\| \leq \sup_x |f(x)|$, if $f \in \mathcal{H}_0$, where $\|f\|$ denotes the norm $(f, f)^{1/2}$.

Let $\mathcal{H}$ be the completion of $\mathcal{H}_0$, then $\mathcal{H}$ is not separable because it has an uncountable orthonormal family $\{e^{i\lambda x}\}_{\lambda\in\mathbb{R}}$.

Closely related to $\mathcal{H}_0$ and $\mathcal{H}$ is the space of **almost periodic (AP)** functions. A continuous function F defined on $\mathbb{R}$ is called AP if it is the uniform limit (on $\mathbb{R}$) of elements in $\mathcal{H}_0$. Such functions can be identified with (certain) elements in the completion $\mathcal{H}$: We have $\mathcal{H}_0 \subseteq AP \subseteq \mathcal{H}$.

Function in AP has an intuitive equivalent definition: For every $\varepsilon > 0$ we can find a length $L = L_\varepsilon$ for a given $F \in AP$ such that any interval $I \subseteq \mathbb{R}$ of length L contains an "almost period" τ satisfying

$$\sup_x |F(x+\tau) - F(x)| < \varepsilon.$$

A third equivalent characterization is that F is in AP iff every sequence $F(x + h_n)$ of translates of F contains a subsequence that converges uniformly.

4.3 Supplementaries

Exercise 4.a (Each Topological Subspace is Separable of a Separable Space)

For readers who are fimilar with the art of topology, you may skip this warm-up exercise, which lays the foundation for Exercise 4.b below.

Suppose $\mathcal{H}$ is a *separable* Hilbert space and $Y \subseteq X$ is any non-trivial subset. Show that Y is separable in the sense that it contains a countable, dense subset. More specifically, there exists a sequence $\{x_n\}_{n=1}^{\infty}$ of Y such that every non-empty open subset of Y contains at least one element of the sequence. We say a subset $U \subseteq Y$ is open if there exists an open set $V \subseteq \mathcal{H}$, such that $U = V \cap Y$.

Solution [Sketch] The collection of finite combinations of elements of a countable dense family with rational coefficients[2] still has a cardinality of $\aleph_0$. Consequently, the ball centered at points in this collection with rational

[2]If the base field is $\mathbb{C}$, take the coefficients whose real and imaginary parts are both $\mathbb{Q}$.

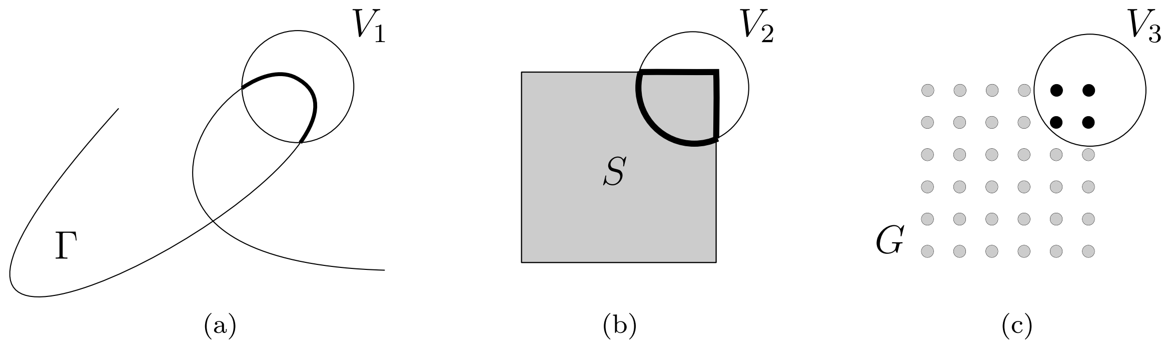

Figure 4.2: Examples of open sets in various subspace of $\mathbb{R}^2$. In (a), Γ is a segment of a curve and V_1 is an open disk, so the bolded part (excluding two end points) is open in Γ. (b) is a somewhat counter-intuitive example. Note the area encircled by the bolded curves is open (with only straight edges included). Nevertheless, this closed cube S is itself open too. Finally, (c) shows a portion of the grid G. Note, however, that any subset of G is open as well. Combined with the Euclidean distance, G is a quintessential discrete metric space.

radius forms the basis of open sets in $\mathcal{H}$. For such a ball B that intersects with Y, choose a point p in $B \cap Y$, and the collection of p is denumerable and separable at the same time. ■

Exercise 4.b (*Saparable* Hilbert Space has only Countable Orthonormal Families)

If a Hilbert space $\mathcal{H}$ is separable, then every orthonormal family of $\mathcal{H}$ is countable.

Solution Of course, $\mathcal{H}$ is separable iff it has one orthonormal basis.

If $\mathcal{H}$ has a orthonormal family $\{e_\lambda\}_{\lambda\in\Lambda}$ that is uncountable, then we take it as the subspace. Note, however, that it has a pairwise distance of 2, so the dense subset of $\{e_\lambda\}_{\lambda\in\Lambda}$ is the set itself. According to Exercise 4.a, it has a denumerable dense subset, contradicting the cardinality of Λ. ■