Measure Theory Notes

Contents

Chapter 2E: Convergence of Measurable Functions

(2.82) Definition: Pointwise and Uniform Convergence

Let X be a set, and let $f_1, f_2, \dots$ be a sequence of functions $f_k : X \to \mathbb{R}$. Let $f : X \to \mathbb{R}$.

Definition 1 (Pointwise convergence)**.** We say $f_k \to f$ *pointwise on* X if for every $x \in X$,

$$\lim_{k\to\infty} f_k(x) = f(x).$$

Definition 2 (Uniform convergence)**.** We say $f_k \to f$ *uniformly on* X if for every $\varepsilon > 0$ there exists $n \in \mathbb{N}$ such that

$$|f_k(x) - f(x)| \le \varepsilon \quad \text{for all } k \ge n \text{ and all } x \in X.$$

(2.83) Example: Pointwise but Not Uniform

Define $f : [-1, 1] \to \mathbb{R}$ by

$$f(x) = \begin{cases} 1, & x \neq 0, \\ 2, & x = 0. \end{cases}$$

Define $f_k : [-1, 1] \to \mathbb{R}$ by

$$f_k(x) = \begin{cases} 1, & |x| \ge \frac{1}{k}, \\ 2 - k|x|, & |x| < \frac{1}{k}. \end{cases}$$

Then $f_k(x) \to f(x)$ pointwise on $[-1, 1]$:

- If $x \neq 0$, choose K so that $1/K < |x|$. For $k \ge K$, we have $|x| \ge 1/k$, hence $f_k(x) = 1 = f(x)$.

- If $x = 0$, then $f_k(0) = 2 = f(0)$ for all k.

But the convergence is *not* uniform. Intuitively, each f_k has a narrow "spike" near 0 that never disappears uniformly; more formally, for any k and any very small $x \neq 0$ with $|x| < 1/k$, we have $f(x) = 1$ while $f_k(x)$ can be arbitrarily close to 2, so $\sup_{x\in[-1,1]} |f_k(x) - f(x)|$ does not go to 0.

(2.84) Result: Uniform Limit of Continuous Functions is Continuous

Proposition 1. *Suppose $B \subseteq \mathbb{R}$ and $f_1, f_2, \ldots$ is a sequence of functions $f_k : B \to \mathbb{R}$ that converges uniformly on B to a function $f : B \to \mathbb{R}$. Fix $b \in B$. If each f_k is continuous at b, then f is continuous at b.*

(2.85) Egorov's Theorem

Theorem 1 (Egorov)**.** *Let $(X, \mathcal{S}, \mu)$ be a measure space with $\mu(X) < \infty$. Suppose $f_1, f_2, \ldots$ is a sequence of $\mathcal{S}$-measurable functions $f_k : X \to \mathbb{R}$ that converges pointwise on X to a function $f : X \to \mathbb{R}$. Then for every $\varepsilon > 0$, there exists a set $E \in \mathcal{S}$ such that*

$$\mu(X \setminus E) \leq \varepsilon \quad \textit{and} \quad f_k \to f \textit{ uniformly on } E.$$

(2.88) Definition: Simple Functions

Definition 3. A function is called *simple* if it takes on only finitely many values.

(2.89) Result: Approximation by Simple Functions

Proposition 2 (Approximation by simple functions)**.** *Suppose $(X, \mathcal{S})$ is a measure space and $f : X \to [0, \infty]$ is $\mathcal{S}$-measurable. Then there exists a sequence $f_1, f_2, \ldots$ of functions $X \to \mathbb{R}$ such that:*

1. *Each f_k is a simple, $\mathcal{S}$-measurable function.*
2. *$f_k(x) \leq f_{k+1}(x) \leq f(x)$ for all $k \in \mathbb{N}$ and all $x \in X$.*
3. *$\lim\limits_{k\to\infty} f_k(x) = f(x)$ for every $x \in X$.*
4. *$f_k \to f$ uniformly on X* iff *f is bounded.*

(2.91) Lusin's Theorem

Theorem 2 (Lusin)**.** *Suppose $g : \mathbb{R} \to \mathbb{R}$ is a Borel measurable function. Then for every $\varepsilon > 0$, there exists a closed set $F \subset \mathbb{R}$ such that*

$$m(\mathbb{R} \setminus F) < \varepsilon \quad \textit{and} \quad g|_F \textit{ is continuous on } F,$$

where m denotes Lebesgue measure.

(2.92) Result: Continuous Extension of Continuous Functions

Proposition 3 (Continuous extension from a closed set)**.** *Every continuous function on a closed subset of $\mathbb{R}$ can be extended to a continuous function on all of $\mathbb{R}$.*

More precisely: if $F \subset \mathbb{R}$ is closed and $g : F \to \mathbb{R}$ is continuous, then there exists a continuous function $h : \mathbb{R} \to \mathbb{R}$ such that $h|_F = g$.

(2.93) Lusin's Theorem: Second Version (on a Set E)

Theorem 3 (Lusin, restricted to a measurable set)**.** *Suppose $E \subset \mathbb{R}$ and $g : E \to \mathbb{R}$ is a Borel measurable function. Then for every $\varepsilon > 0$, there exist*

- *a closed set $F \subset E$, and*
- *a continuous function $h : \mathbb{R} \to \mathbb{R}$,*

such that

$$m(E \setminus F) \le \varepsilon \quad \text{and} \quad h|_F = g|_F,$$

where m denotes Lebesgue measure.

Interpretation: a Borel measurable function can be modified on a set of arbitrarily small Lebesgue measure to agree with a continuous function (on a large closed subset of its domain).

(2.94) Definition: Lebesgue Measurable Function

Definition 4 (Lebesgue measurability)**.** Let $A \subset \mathbb{R}$. A function $f : A \to \mathbb{R}$ is called *Lebesgue measurable* if $f^{-1}(B)$ is a Lebesgue measurable subset of $\mathbb{R}$ for every Borel set $B \subset \mathbb{R}$.

- If $f : A \to \mathbb{R}$ is Lebesgue measurable, then A is Lebesgue measurable since $A = f^{-1}(\mathbb{R})$.
- If A is Lebesgue measurable and $\mathcal{S}$ denotes the σ-algebra of all Lebesgue measurable subsets of A, then the definition above is the standard definition of $\mathcal{S}$-measurability.

(2.95) Result: Lebesgue Measurable ⇒ Almost Borel Measurable

Proposition 4 (Equal a.e. to a Borel measurable function)**.** *Suppose $f : \mathbb{R} \to \mathbb{R}$ is Lebesgue measurable. Then there exists a Borel measurable function $g : \mathbb{R} \to \mathbb{R}$ such that*

$$m\big(\{x \in \mathbb{R} : \ g(x) \ne f(x)\}\big) = 0.$$

Exercises 2E (selected)

(2) **If X is finite, pointwise convergence implies uniform convergence.**

Suppose X is finite and $f_n : X \to \mathbb{R}$ with $f_n(x) \to f(x)$ for each $x \in X$. Fix $\varepsilon > 0$. For each $x \in X$, choose $N_x \in \mathbb{N}$ such that

$$|f_n(x) - f(x)| \le \varepsilon \quad \text{for all } n \ge N_x.$$

Since X is finite, $N := \max_{x \in X} N_x$ exists. Then for all $n \ge N$ and all $x \in X$,

$$|f_n(x) - f(x)| \le \varepsilon,$$

so $f_n \to f$ uniformly on X.

(3) **Continuous $f_n : [0,1] \to \mathbb{R}$ converging pointwise to an unbounded function.**

For $n \in \mathbb{N}$, define $f_n : [0,1] \to \mathbb{R}$ by

$$f_n(x) = \begin{cases} \dfrac{1}{x - \frac{1}{2}}, & |x - \frac{1}{2}| \ge \dfrac{1}{n}, \\ n^2\big(x - \frac{1}{2}\big), & |x - \frac{1}{2}| < \dfrac{1}{n}. \end{cases}$$

At the junction points $x = \frac{1}{2} \pm \frac{1}{n}$, both formulas give $\pm n$, so f_n is continuous on $[0,1]$. For $x \neq \frac{1}{2}$, eventually $|x - \frac{1}{2}| \geq 1/n$, hence $f_n(x) = \frac{1}{x-\frac{1}{2}}$ for all large n. Also $f_n(\frac{1}{2}) = 0$ for all n. Therefore $f_n \to f$ pointwise, where

$$f(x) = \begin{cases} \dfrac{1}{x - \frac{1}{2}}, & x \neq \frac{1}{2}, \\ 0, & x = \frac{1}{2}, \end{cases}$$

and f is not bounded on $[0,1]$ (it blows up near $x = \frac{1}{2}$).

(5) **Egorov can fail if $\mu(X) = \infty$.**

On $(\mathbb{R}, \mathcal{L}, m)$ with Lebesgue measure $m(\mathbb{R}) = \infty$, let

$$f_n = \mathbf{1}_{(-n,n)}.$$

Then $f_n(x) \to 1$ for every $x \in \mathbb{R}$, so the pointwise (hence a.e.) limit is $f \equiv 1$.

Take any measurable $E \subset \mathbb{R}$ with $m(\mathbb{R} \setminus E) < \varepsilon$. Then E must be unbounded, so for each n there exists $x \in E$ with $|x| \geq n$, hence $f_n(x) = 0$ while $f(x) = 1$. Therefore

$$\sup_{x \in E} |f_n(x) - f(x)| = 1 \quad \text{for all } n,$$

so $f_n \not\to f$ uniformly on E. Thus the "almost uniform" conclusion of Egorov's theorem can fail without the hypothesis $m(X) < \infty$.

(6) **If $f_n(x) \to \infty$ pointwise and $\mu(X) < \infty$, then $f_n \to \infty$ almost uniformly.**

Let $(X, \mathcal{S}, \mu)$ satisfy $\mu(X) < \infty$, and let $f_1, f_2, \ldots$ be $\mathcal{S}$-measurable with $\lim_{n\to\infty} f_n(x) = \infty$ for each $x \in X$. Fix $\varepsilon > 0$.

For each $m \in \mathbb{N}$, define the truncations

$$f_n^{(m)}(x) := \min\{f_n(x),\, m\}.$$

Each $f_n^{(m)}$ is measurable and bounded, and since $f_n(x) \to \infty$, we have $f_n^{(m)}(x) \to m$ pointwise.

Apply Egorov's theorem (using $\mu(X) < \infty$) to the sequence $\{f_n^{(m)}\}_{n=1}^{\infty}$ for each fixed m: there exists $E_m \in \mathcal{S}$ such that

$$\mu(X \setminus E_m) \leq \frac{\varepsilon}{2^m} \quad \text{and} \quad f_n^{(m)} \to m \text{ uniformly on } E_m.$$

Hence there exists N_m such that for all $n \geq N_m$,

$$\sup_{x \in E_m} \left| f_n^{(m)}(x) - m \right| < \frac{1}{2},$$

so for all $n \geq N_m$ and $x \in E_m$, we have $f_n^{(m)}(x) > m - \frac{1}{2}$, which implies $f_n(x) > m - \frac{1}{2}$.

Now set

$$E := \bigcap_{m=1}^{\infty} E_m.$$

Then

$$\mu(X \setminus E) \le \sum_{m=1}^{\infty} \mu(X \setminus E_m) \le \sum_{m=1}^{\infty} \frac{\varepsilon}{2^m} = \varepsilon.$$

Finally, to see uniform divergence on E: given $T > 0$, choose m with $m - \frac{1}{2} \ge T$ and let $N := N_m$. For all $n \ge N$ and all $x \in E \subset E_m$, we get

$$f_n(x) > m - \tfrac{1}{2} \ge T.$$

Thus for every $T > 0$ there exists N such that $n \ge N \Rightarrow f_n(x) \ge T$ for all $x \in E$, i.e. $f_n \to \infty$ uniformly on E.

Exercises 2E (continued)

(8) **An "Egorov"-type statement on $\big(\mathbb{Z}^+, \mathcal{P}(\mathbb{Z}^+), \mu\big)$ with $\mu(E) = \sum_{n\in E} 2^{-n}$.**

Let μ be the measure on $(\mathbb{Z}^+, \mathcal{P}(\mathbb{Z}^+))$ defined by

$$\mu(E) := \sum_{n \in E} 2^{-n}.$$

Claim: for every $\varepsilon > 0$, there exists a set $E \subset \mathbb{Z}^+$ with $\mu(\mathbb{Z}^+ \setminus E) \le \varepsilon$ such that *every* pointwise convergent sequence $f_1, f_2, \dots : \mathbb{Z}^+ \to \mathbb{R}$ converges uniformly on E.

Proof. Fix $\varepsilon > 0$. Choose $N \in \mathbb{Z}^+$ such that $2^{-N} < \varepsilon$, and set

$$E := \{1, 2, \dots, N\}.$$

Then

$$\mu(\mathbb{Z}^+ \setminus E) = \sum_{n > N} 2^{-n} = 2^{-N} < \varepsilon.$$

Now let $f_k : \mathbb{Z}^+ \to \mathbb{R}$ be any sequence with $f_k \to f$ pointwise on $\mathbb{Z}^+$. Since E is finite, given $\delta > 0$ and each $m \in E$ there exists K_m such that

$$|f_k(m) - f(m)| < \delta \quad \text{for all } k \ge K_m.$$

Let $K := \max_{m\in E} K_m$. Then for all $k \ge K$ and all $m \in E$,

$$|f_k(m) - f(m)| < \delta,$$

i.e. $\sup_{m\in E} |f_k(m) - f(m)| < \delta$, so $f_k \to f$ uniformly on E. □

(9) **Continuity on a finite union of disjoint closed sets.**

Let $F_1, \dots, F_n$ be pairwise disjoint closed subsets of $\mathbb{R}$, and set $U := \bigcup_{k=1}^{n} F_k$. Suppose $g : U \to \mathbb{R}$ satisfies that $g|_{F_k}$ is continuous on F_k for each $k \in \{1, \dots, n\}$. Then g is continuous on U (with the subspace topology).

Proof. Fix $x \in U$, and choose m such that $x \in F_m$. For each $k \ne m$, since F_k is closed and $x \notin F_k$, we have $\operatorname{dist}(x, F_k) > 0$. Because there are only finitely many such k, the minimum

$$\delta_0 := \min_{k \ne m} \operatorname{dist}(x, F_k) > 0$$

is positive. Hence

$$(x-\delta_0/2,\ x+\delta_0/2)\cap U\subset F_m.$$

Now use continuity of $g|_{F_m}$ at x: given $\varepsilon>0$, choose $\delta_1>0$ such that

$$y\in F_m,\ |y-x|<\delta_1 \implies |g(y)-g(x)|<\varepsilon.$$

Let $\delta:=\min\{\delta_0/2,\delta_1\}$. If $y\in U$ and $|y-x|<\delta$, then $y\in F_m$ and therefore $|g(y)-g(x)|<\varepsilon$. Thus g is continuous at x, and since x was arbitrary, g is continuous on U. □

(14) **A function finite on a set of infinite measure.**

Suppose $b_1,b_2,\dots$ is a sequence of real numbers. Define $f:\mathbb{R}\to[0,\infty]$ by

$$f(x):=\begin{cases}\displaystyle\sum_{k=1}^{\infty}\frac{1}{4^k|x-b_k|}, & x\notin\{b_1,b_2,\dots\},\\ \infty, & x\in\{b_1,b_2,\dots\}.\end{cases}$$

Prove that the set $\{x\in\mathbb{R}: f(x)\le 1\}$ has infinite Lebesgue measure.

Proof. Fix a constant $c>1$ (e.g. $c=2$) and set $r_k:=c\,2^{-k}$. Let

$$U:=\bigcup_{k=1}^{\infty}(b_k-r_k,\ b_k+r_k).$$

Then

$$|U|\le\sum_{k=1}^{\infty}|(b_k-r_k,\ b_k+r_k)|=\sum_{k=1}^{\infty}2r_k=2c\sum_{k=1}^{\infty}2^{-k}=2c<\infty,$$

so $|\mathbb{R}\setminus U|=\infty$.

If $x\in\mathbb{R}\setminus U$, then $|x-b_k|\ge r_k$ for every k, hence

$$f(x)\le\sum_{k=1}^{\infty}\frac{1}{4^k r_k}=\sum_{k=1}^{\infty}\frac{1}{4^k\cdot c\,2^{-k}}=\frac{1}{c}\sum_{k=1}^{\infty}2^{-k}=\frac{1}{c}\le 1.$$

Therefore $\mathbb{R}\setminus U\subset\{x: f(x)\le 1\}$, and since $|\mathbb{R}\setminus U|=\infty$, we conclude

$$\big|\{x\in\mathbb{R}: f(x)\le 1\}\big|=\infty.$$

□

(15) **Lebesgue measurable on a Borel set is a.e. equal to a Borel measurable function.**

Suppose B is a Borel set and $f:B\to\mathbb{R}$ is Lebesgue measurable. Show there exists a Borel measurable function $g:B\to\mathbb{R}$ such that

$$m\big(\{x\in B:\ g(x)\ne f(x)\}\big)=0.$$

Proof sketch (regularity construction). For each $q\in\mathbb{Q}$, set $A_q:=\{x\in B:\ f(x)<q\}$. Each A_q is Lebesgue measurable (as a subset of $\mathbb{R}$). By regularity of Lebesgue measure, choose a Borel set $E_q\subset\mathbb{R}$ such that

$$A_q\subset E_q \quad\text{and}\quad m(E_q\setminus A_q)=0.$$

Replace E_q by
$$\widetilde{E}_q := \bigcap_{\substack{r\in\mathbb{Q}\\ r>q}} E_r$$
to ensure monotonicity ($q < r \Rightarrow \widetilde{E}_q \subset \widetilde{E}_r$); still $m(\widetilde{E}_q \triangle A_q) = 0$.
Define $g : B \to \mathbb{R}$ by
$$g(x) := \inf\{q \in \mathbb{Q} :\ x \in \widetilde{E}_q \cap B\}.$$
Then for any $\alpha \in \mathbb{R}$,
$$\{x \in B :\ g(x) < \alpha\} = \bigcup_{\substack{q\in\mathbb{Q}\\ q<\alpha}} (\widetilde{E}_q \cap B),$$
which is a Borel subset of B, so g is Borel measurable.
Let $N := \bigcup_{q\in\mathbb{Q}}(\widetilde{E}_q \triangle A_q)$. Then $m(N) = 0$. If $x \in B \setminus N$, then for every $q \in \mathbb{Q}$,
$$x \in \widetilde{E}_q \iff x \in A_q \iff f(x) < q,$$
hence $g(x) = \inf\{q \in \mathbb{Q} :\ f(x) < q\} = f(x)$. Therefore $g = f$ a.e. on B. $\square$

Chapter 3A: Integration with Respect to a Measure

Definition: $\mathcal{S}$-partition

Let $\mathcal{S}$ be a σ-algebra on a set X. An *$\mathcal{S}$-partition* of X is a finite collection $A_1, \ldots, A_m \in \mathcal{S}$ of pairwise disjoint sets such that
$$A_1 \cup \cdots \cup A_m = X.$$

Definition: Lower Lebesgue sum

Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $f : X \to [0, \infty]$ is $\mathcal{S}$-measurable, and $P = \{A_1, \ldots, A_m\}$ is an $\mathcal{S}$-partition of X. The *lower Lebesgue sum* is
$$L(f, P) := \sum_{j=1}^{m} \mu(A_j) \inf_{x\in A_j} f(x).$$

Definition: Integral of a nonnegative measurable function

If $(X, \mathcal{S}, \mu)$ is a measure space and $f : X \to [0, \infty]$ is $\mathcal{S}$-measurable, define
$$\int f\, d\mu \ := \ \sup\{L(f, P) : P \text{ is an } \mathcal{S}\text{-partition of } X\}.$$

Interpretation via simple-function approximation

For a partition $P = \{A_1, \ldots, A_m\}$, define the (simple) step function
$$s_P(x) := \sum_{j=1}^{m} \big(\inf_{A_j} f\big)\, \chi_{A_j}(x).$$

Then s_P is $\mathcal{S}$-measurable, $0 \le s_P \le f$, and

$$\int s_P \, d\mu = \sum_{j=1}^{m} \mu(A_j) \inf_{A_j} f = L(f, P).$$

Taking the supremum over partitions corresponds to taking the best approximation from below.

(3.4) Definition/Fact: Integral of a characteristic function

If $E \in \mathcal{S}$, then

$$\int \chi_E \, d\mu = \mu(E).$$

Example: Counting measure gives summation

If μ is counting measure on $\mathbb{Z}^+$ and $b_1, b_2, \dots \ge 0$, view b as a function $b : \mathbb{Z}^+ \to [0, \infty)$ with $b(k) = b_k$. Then

$$\int b \, d\mu = \sum_{k=1}^{\infty} b_k.$$

(3.7) Result: Integral of a simple function

If $(X, \mathcal{S}, \mu)$ is a measure space, $E_1, \dots, E_n \in \mathcal{S}$ are disjoint, and $c_1, \dots, c_n \in [0, \infty)$, then

$$\int \left(\sum_{k=1}^{n} c_k \chi_{E_k} \right) d\mu = \sum_{k=1}^{n} c_k \, \mu(E_k).$$

(Proof idea: one direction uses the partition $\{E_1, \dots, E_n, X \setminus \cup_k E_k\}$; the other direction compares $L(\cdot, P)$ for an arbitrary partition P to the weighted sum $\sum_k c_k \mu(E_k)$.)

(3.8) Result: Integration is order-preserving

If $f, g : X \to [0, \infty]$ are $\mathcal{S}$-measurable and $f(x) \le g(x)$ for all $x \in X$, then

$$\int f \, d\mu \le \int g \, d\mu.$$

(3.9) Restatement: Supremum over simple functions dominated by f

If $f : X \to [0, \infty]$ is $\mathcal{S}$-measurable, then

$$\int f \, d\mu = \sup \left\{ \sum_{j=1}^{m} c_j \, \mu(A_j) \ : \ \begin{array}{l} A_1, \dots, A_m \in \mathcal{S} \text{ disjoint}, \ c_1, \dots, c_m \in [0, \infty), \\ \sum_{j=1}^{m} c_j \, \chi_{A_j}(x) \le f(x) \text{ for all } x \in X \end{array} \right\}.$$

(Used later for the Monotone Convergence Theorem.)

(3.11) Result: Monotone Convergence Theorem (MCT)

Theorem 4 (Monotone Convergence)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $0 \le f_1 \le f_2 \le \cdots$ is an increasing sequence of $\mathcal{S}$-measurable functions. Define $f : X \to [0, \infty]$ by*

$$f(x) := \lim_{k\to\infty} f_k(x).$$

Then

$$\lim_{k\to\infty} \int f_k \, d\mu \;=\; \int f \, d\mu.$$

Proof (standard). Since $f_k \le f$, we have $\int f_k \, d\mu \le \int f \, d\mu$ for all k, hence $\lim_{k\to\infty} \int f_k \, d\mu \le \int f \, d\mu$.

For the reverse inequality, let $s = \sum_{j=1}^{m} c_j \chi_{A_j}$ be any nonnegative simple function with $s \le f$. Fix $t \in (0, 1)$ and define

$$E_k := \{x \in X : \; f_k(x) \ge t\, s(x)\}.$$

Then $E_k \in \mathcal{S}$, $E_k \uparrow X$, and for each k we have $f_k \ge t\, s\, \chi_{E_k}$. Therefore,

$$\int f_k \, d\mu \;\ge\; t \int s\, \chi_{E_k} \, d\mu \;=\; t \sum_{j=1}^{m} c_j \, \mu(A_j \cap E_k).$$

Letting $k \to \infty$ and using $\mu(A_j \cap E_k) \uparrow \mu(A_j)$ gives

$$\lim_{k\to\infty} \int f_k \, d\mu \;\ge\; t \sum_{j=1}^{m} c_j \, \mu(A_j) \;=\; t \int s \, d\mu.$$

Now let $t \uparrow 1$ to obtain $\lim_{k\to\infty} \int f_k \, d\mu \ge \int s \, d\mu$. Finally, take the supremum over all such simple $s \le f$ (by the definition/restatement of $\int f \, d\mu$) to get $\lim_{k\to\infty} \int f_k \, d\mu \ge \int f \, d\mu$. □

(3.13) Result: "Integral-type" sums for simple functions

Proposition 5 (Consistency across representations)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space. If*

$$\sum_{j=1}^{m} a_j \chi_{A_j} \;=\; \sum_{k=1}^{n} b_k \chi_{B_k} \quad \textit{(pointwise on } X\textit{)},$$

where $A_1, \ldots, A_m, B_1, \ldots, B_n \in \mathcal{S}$ and $a_j, b_k \in [0, \infty)$, then

$$\sum_{j=1}^{m} a_j \, \mu(A_j) \;=\; \sum_{k=1}^{n} b_k \, \mu(B_k).$$

Remark. A representation of a simple function $h : X \to [0, \infty)$ as $\sum_{k=1}^{n} c_k \chi_{E_k}$ is not unique. Imposing that the coefficients c_k are distinct and that the sets E_k are nonempty, pairwise disjoint, and satisfy $E_1 \cup \cdots \cup E_n = X$ yields the *standard representation.* The proposition says every representation yields the same "integral-type" sum, so the integral of a simple function is well-defined.

(3.15) Result: Integral of a linear combination of characteristic functions

If $E_1, \dots, E_n \in \mathcal{S}$ are pairwise disjoint and $c_1, \dots, c_n \in [0, \infty)$, then

$$\int \Big(\sum_{k=1}^{n} c_k \chi_{E_k} \Big)\, d\mu \;=\; \sum_{k=1}^{n} c_k\, \mu(E_k).$$

(If the E_k are not disjoint, first refine to a disjoint partition by taking intersections of E_k and complements.)

(3.16) Result: Additivity of integration (nonnegative functions)

Proposition 6 (Additivity for $f, g \geq 0$)**.** *Let $(X, \mathcal{S}, \mu)$ be a measure space and let $f, g : X \to [0, \infty]$ be $\mathcal{S}$-measurable. Then*

$$\int (f+g)\, d\mu \;=\; \int f\, d\mu + \int g\, d\mu.$$

(3.17) Definition: Positive and negative parts

Suppose $f : X \to [-\infty, \infty]$ is a function. Define $f^+, f^- : X \to [0, \infty]$ by

$$f^+(x) := \max\{f(x), 0\}, \qquad f^-(x) := \max\{-f(x), 0\}.$$

Then

$$f = f^+ - f^-, \qquad |f| = f^+ + f^-.$$

(3.18) Definition: Integral of a real-valued (extended) measurable function

Let $(X, \mathcal{S}, \mu)$ be a measure space and let $f : X \to [-\infty, \infty]$ be $\mathcal{S}$-measurable. Assume at least one of $\int f^+\, d\mu$ or $\int f^-\, d\mu$ is finite (equivalently, not both are $+\infty$). Define

$$\int f\, d\mu \;:=\; \int f^+\, d\mu \;-\; \int f^-\, d\mu.$$

Notes:

- If $f \geq 0$, then $f^- = 0$ and this reduces to the nonnegative definition.
- $\int |f|\, d\mu < \infty \iff \int f^+\, d\mu < \infty$ and $\int f^-\, d\mu < \infty$.

(3.20) Result: Integration is homogeneous

If $\int f\, d\mu$ is defined and $c \in \mathbb{R}$, then (when the RHS is defined)

$$\int (cf)\, d\mu = c \int f\, d\mu.$$

(For $c \geq 0$ this is immediate from $(cf)^{\pm} = cf^{\pm}$; extend to $c < 0$ using $(cf)^+ = (-c)f^-$ and $(cf)^- = (-c)f^+$.)

(3.21) Result: Additivity for integrable functions

If f, g are $\mathcal{S}$-measurable and $\int |f|\, d\mu < \infty$, $\int |g|\, d\mu < \infty$, then

$$\int (f+g)\, d\mu = \int f\, d\mu + \int g\, d\mu.$$

(3.22) Result: Integration is order-preserving

If f, g are $\mathcal{S}$-measurable, $f \leq g$ pointwise, and $\int f\,d\mu, \int g\,d\mu$ are defined (e.g. both integrable), then

$$\int f\,d\mu \leq \int g\,d\mu.$$

(3.23) Result: Absolute value inequality

If $\int f\,d\mu$ is defined, then

$$\left|\int f\,d\mu\right| \leq \int |f|\,d\mu.$$

(Proof idea: $\int f = \int f^+ - \int f^-$ and $\int |f| = \int f^+ + \int f^-$.)

Exercises 3A (selected)

(2) **Dirac measure.** Suppose X is a set, $\mathcal{S}$ is a σ-algebra on X, and $c \in X$. Define the Dirac measure δ_c on $(X, \mathcal{S})$ by

$$\delta_c(E) := \begin{cases} 1, & c \in E, \\ 0, & c \notin E, \end{cases} \qquad E \in \mathcal{S}$$

(note: $\{c\}$ need not be in $\mathcal{S}$). Prove that if $f : X \to [0, \infty]$ is $\mathcal{S}$-measurable, then

$$\int f\,d\delta_c = f(c).$$

Solution. First, for an indicator χ_E we have

$$\int \chi_E\,d\delta_c = \delta_c(E) = \chi_E(c).$$

Next, for a simple function $s = \sum_{k=1}^n a_k \chi_{E_k}$ (with $a_k \geq 0$) we get

$$\int s\,d\delta_c = \sum_{k=1}^n a_k\,\delta_c(E_k) = \sum_{k=1}^n a_k\,\chi_{E_k}(c) = s(c).$$

For general measurable $f \geq 0$, choose simple $s_n \uparrow f$ pointwise. Then by MCT,

$$\int f\,d\delta_c = \lim_{n\to\infty} \int s_n\,d\delta_c = \lim_{n\to\infty} s_n(c) = f(c).$$

□

(5) **Counting measure gives summation.** Let $(X, \mathcal{A}, \nu)$ be a measure space where ν is counting measure:

$$\nu(E) = \#E \in \{0, 1, 2, \ldots, \infty\}.$$

Verify that integration w.r.t. ν equals summation.

Solution. (a) If $E \in \mathcal{A}$, then

$$\int \chi_E\,d\nu = \nu(E) = \sum_{x \in X} \chi_E(x).$$

If $s = \sum_{k=1}^{n} a_k \chi_{E_k}$ is nonnegative simple with disjoint E_k, then

$$\int s\, d\nu = \sum_{k=1}^{n} a_k\, \nu(E_k) = \sum_{k=1}^{n} a_k \sum_{x\in X} \chi_{E_k}(x) = \sum_{x\in X} s(x).$$

(b) For $f : X \to [0, \infty]$ measurable, for each finite $F \subset X$ define $s_F := f\, \chi_F$. Then s_F is simple, $0 \le s_F \le f$, and as F increases over finite subsets, $s_F \uparrow f$ pointwise. Hence by MCT,

$$\int f\, d\nu = \sup_{F\subset X,\ F \text{ finite}} \int s_F\, d\nu = \sup_{F\subset X,\ F \text{ finite}} \sum_{x\in F} f(x) = \sum_{x\in X} f(x),$$

where the infinite sum is defined as the supremum over finite partial sums.

(c) If f is signed with $\int |f|\, d\nu < \infty$, then write $f = f^+ - f^-$ and apply the nonnegative case to get

$$\int f\, d\nu = \sum_{x\in X} f(x),$$

with absolute convergence. $\square$

(7) **Weighted counting measure.** Suppose $\mathcal{S} = \mathcal{P}(X)$ (all subsets), and $w : X \to [0, \infty]$. Define μ on $(X, \mathcal{S})$ by

$$\mu(E) := \sum_{x\in E} w(x),$$

where the infinite sum is defined as $\sup\{\sum_{x\in F} w(x) : F \subset E,\ F \text{ finite}\}$. Prove that for any $f : X \to [0, \infty]$,

$$\int f\, d\mu \;=\; \sum_{x\in X} w(x) f(x),$$

with the RHS again understood as a supremum over finite partial sums.

Solution. (a) For $E \subset X$,

$$\int \chi_E\, d\mu = \mu(E) = \sum_{x\in E} w(x) = \sum_{x\in X} w(x)\chi_E(x).$$

(b) If $s = \sum_{j=1}^{m} a_j \chi_{E_j}$ is simple with disjoint E_j, then using the simple-function formula,

$$\int s\, d\mu = \sum_{j=1}^{m} a_j\, \mu(E_j) = \sum_{j=1}^{m} a_j \sum_{x\in E_j} w(x) = \sum_{x\in X} w(x) s(x).$$

(c) For general $f \ge 0$, choose simple $s_n \uparrow f$. Then by MCT,

$$\int f\, d\mu = \lim_{n\to\infty} \int s_n\, d\mu = \lim_{n\to\infty} \sum_{x\in X} w(x) s_n(x).$$

Since $w(x)s_n(x) \uparrow w(x)f(x)$ pointwise in n, the RHS equals $\sum_{x\in X} w(x)f(x)$ (as a supremum over finite partial sums). $\square$

Exercises 3A (continued)

10. Suppose $(X,\mathcal{S},\mu)$ is a measure space and $f_1,f_2,\ldots$ is a sequence of nonnegative $\mathcal{S}$-measurable functions. Define $f : X \to [0,\infty]$ by

$$f(x) := \sum_{k=1}^{\infty} f_k(x).$$

Prove that

$$\int f\,d\mu \;=\; \sum_{k=1}^{\infty} \int f_k\,d\mu.$$

Proof. Let the partial sums be

$$s_n(x) := \sum_{k=1}^{n} f_k(x), \qquad n \in \mathbb{N}.$$

Then each s_n is nonnegative and measurable, and $s_n(x) \uparrow f(x)$ pointwise. By the Monotone Convergence Theorem,

$$\int f\,d\mu \;=\; \lim_{n\to\infty} \int s_n\,d\mu.$$

Since each s_n is a *finite* sum of nonnegative measurable functions,

$$\int s_n\,d\mu \;=\; \sum_{k=1}^{n} \int f_k\,d\mu.$$

Taking $n \to \infty$ gives

$$\int f\,d\mu = \lim_{n\to\infty} \sum_{k=1}^{n} \int f_k\,d\mu = \sum_{k=1}^{\infty} \int f_k\,d\mu.$$

$\square$

11. Suppose $(X,\mathcal{S},\mu)$ is a measure space and $f_1,f_2,\ldots$ are $\mathcal{S}$-measurable functions $X \to \mathbb{R}$ such that

$$\sum_{k=1}^{\infty} \int |f_k|\,d\mu < \infty.$$

Prove there exists $E \in \mathcal{S}$ such that $\mu(X \setminus E) = 0$ and $\lim_{k\to\infty} f_k(x) = 0$ for every $x \in E$.

Proof. Define the nonnegative measurable partial sums

$$g_n(x) := \sum_{k=1}^{n} |f_k(x)|, \qquad n \in \mathbb{N},$$

and set

$$g(x) := \sum_{k=1}^{\infty} |f_k(x)| = \sup_n g_n(x).$$

Then $g_n \uparrow g$ pointwise, so by the Monotone Convergence Theorem,

$$\int g\,d\mu = \lim_{n\to\infty} \int g_n\,d\mu = \lim_{n\to\infty} \sum_{k=1}^{n} \int |f_k|\,d\mu = \sum_{k=1}^{\infty} \int |f_k|\,d\mu < \infty.$$

Hence $g(x) < \infty$ for μ-a.e. x (otherwise $\mu(\{g = \infty\}) > 0$ would force $\int g\,d\mu = \infty$). Let

$$E := \{x \in X : g(x) < \infty\} \in \mathcal{S}.$$

Then $\mu(X \setminus E) = 0$. For each $x \in E$ we have $\sum_{k=1}^{\infty} |f_k(x)| < \infty$, so in particular $|f_k(x)| \to 0$, hence $f_k(x) \to 0$. $\square$

20. Suppose $(X, \mathcal{S}, \mu)$ is a measure space, and $f_1, f_2, \ldots$ is a monotone sequence of $\mathcal{S}$-measurable functions. Define $f : X \to [-\infty, \infty]$ by

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

Prove that if $\int |f_1|\,d\mu < \infty$, then

$$\lim_{n\to\infty} \int f_n\,d\mu \;=\; \int f\,d\mu,$$

(with integrals understood in the extended-real sense).

Proof. First note that $f_1 \in L^1(\mu)$ implies f_1 is finite a.e. and $\int f_1\,d\mu$ is finite. Also, if $h \in L^1(\mu)$ and $p \geq 0$ is measurable, then (in extended-real arithmetic)

$$\int (h + p)\,d\mu = \int h\,d\mu + \int p\,d\mu,$$

and similarly $\int (h - p)\,d\mu = \int h\,d\mu - \int p\,d\mu$ (possibly $-\infty$), since $\int h$ is finite.

Case 1: $f_n \uparrow$. Then $f_n \geq f_1$ pointwise, so define

$$g_n := f_n - f_1 \geq 0, \qquad g := f - f_1 \geq 0.$$

We have $g_n \uparrow g$ pointwise, hence by MCT,

$$\int g_n\,d\mu \to \int g\,d\mu.$$

Using additivity with $h = f_1$ and $p = g_n$,

$$\int f_n\,d\mu = \int (f_1 + g_n)\,d\mu = \int f_1\,d\mu + \int g_n\,d\mu \xrightarrow[n\to\infty]{} \int f_1\,d\mu + \int g\,d\mu = \int (f_1 + g)\,d\mu = \int f\,d\mu.$$

Case 2: $f_n \downarrow$. Then $f_n \leq f_1$ pointwise, so define

$$g_n := f_1 - f_n \geq 0, \qquad g := f_1 - f \geq 0.$$

Again $g_n \uparrow g$, so by MCT, $\int g_n\,d\mu \to \int g\,d\mu$. Using additivity with $h = f_1$ and $p = g_n$,

$$\int f_n\,d\mu = \int (f_1 - g_n)\,d\mu = \int f_1\,d\mu - \int g_n\,d\mu \xrightarrow[n\to\infty]{} \int f_1\,d\mu - \int g\,d\mu = \int (f_1 - g)\,d\mu = \int f\,d\mu.$$

In either monotone case, $\displaystyle\lim_{n\to\infty} \int f_n\,d\mu = \int f\,d\mu.$ $\square$

Chapter 3B: Limits of Integrals and Integrals of Limits

(3.24) Definition (Integration on a subset). Let $(X, \mathcal{S}, \mu)$ be a measure space and let $E \in \mathcal{S}$. If $f : X \to [-\infty, \infty]$ is $\mathcal{S}$-measurable, define

$$\int_E f\, d\mu \;:=\; \int_X \mathbf{1}_E\, f\, d\mu$$

provided the right-hand side is defined (otherwise $\int_E f\, d\mu$ is undefined).

(3.25) Result (Bounding an integral). Let $(X, \mathcal{S}, \mu)$ be a measure space, $E \in \mathcal{S}$, and $f : X \to [-\infty, \infty]$ a function such that $\int_E f\, d\mu$ is defined. Then

$$\Big|\int_E f\, d\mu\Big| \;\le\; \mu(E) \sup_{x\in E} |f(x)|.$$

(Interpret $\sup_{x\in E} |f(x)|$ as $+\infty$ if f is unbounded on E.)

Proof. Let $c := \sup_{x\in E} |f(x)|$. Then $|\mathbf{1}_E f| \le c\, \mathbf{1}_E$, so

$$\Big|\int_E f\, d\mu\Big| = \Big|\int_X \mathbf{1}_E f\, d\mu\Big| \le \int_X |\mathbf{1}_E f|\, d\mu \le \int_X c\, \mathbf{1}_E\, d\mu = c\, \mu(E).$$

□

(3.26) Result (Bounded Convergence Theorem). Let $(X, \mathcal{S}, \mu)$ be a measure space with $\mu(X) < \infty$. Suppose $f_1, f_2, \ldots$ are $\mathcal{S}$-measurable functions $X \to \mathbb{R}$ such that $f_k(x) \to f(x)$ pointwise on X. If there exists $c \in (0, \infty)$ such that

$$|f_k(x)| \le c \quad \text{for all } k \in \mathbb{Z}^+ \text{ and all } x \in X,$$

then

$$\lim_{k\to\infty} \int_X f_k\, d\mu \;=\; \int_X f\, d\mu.$$

Proof (via Egorov). Fix $\varepsilon > 0$. By Egorov's theorem, there exists $E \in \mathcal{S}$ such that

$$\mu(X \setminus E) < \delta \quad \text{and} \quad f_k \to f \text{ uniformly on } E,$$

where $\delta > 0$ will be chosen momentarily. Then

$$\Big|\int_X f_k\, d\mu - \int_X f\, d\mu\Big| \le \int_{X\setminus E} |f_k|\, d\mu + \int_{X\setminus E} |f|\, d\mu + \int_E |f_k - f|\, d\mu.$$

Since $|f_k| \le c$ and also $|f| \le c$ pointwise (limit of uniformly bounded functions),

$$\int_{X\setminus E} |f_k|\, d\mu + \int_{X\setminus E} |f|\, d\mu \le 2c\, \mu(X \setminus E).$$

Choose δ so that $2c\,\delta < \varepsilon/2$. Uniform convergence on E implies $\sup_{x\in E} |f_k(x) - f(x)| \to 0$, hence for k large,

$$\int_E |f_k - f|\, d\mu \le \mu(E) \sup_{x\in E} |f_k - f| < \varepsilon/2$$

(using $\mu(E) \le \mu(X) < \infty$). Therefore the total difference is $< \varepsilon$ for k large. □

Result (Sets of measure 0 in integration). If $f, g : X \to [-\infty, \infty]$ are $\mathcal{S}$-measurable and

$$\mu(\{x \in X : \ f(x) \neq g(x)\}) = 0,$$

then

$$\int_X f \, d\mu = \int_X g \, d\mu,$$

and likewise $\int_E f \, d\mu = \int_E g \, d\mu$ for every $E \in \mathcal{S}$.

Section 3B: Almost everywhere and dominated convergence

(3.27) Definition (Almost every). Let $(X, \mathcal{S}, \mu)$ be a measure space. A set $E \in \mathcal{S}$ is said to contain μ-*almost every* element of X if $\mu(X \setminus E) = 0$.

(Equivalently: a property holds *a.e.* if it fails only on a set of μ-measure 0.)

Example. Almost every real number is irrational (with respect to Lebesgue measure on $\mathbb{R}$) since $\lambda(\mathbb{Q}) = 0$.

(3.28) Result (Integrals on small sets are small). Let $(X, \mathcal{S}, \mu)$ be a measure space and let $g : X \to [0, \infty]$ be $\mathcal{S}$-measurable with

$$\int_X g \, d\mu < \infty.$$

Then for every $\varepsilon > 0$ there exists $\delta > 0$ such that for every $B \in \mathcal{S}$ with $\mu(B) < \delta$,

$$\int_B g \, d\mu < \varepsilon.$$

(In words: if a nonnegative function has finite integral, then its integral over *all* sufficiently small-measure sets is small.)

(3.29) Result (Integrable functions live mostly on sets of finite measure). Let $(X, \mathcal{S}, \mu)$ be a measure space and let $g : X \to [0, \infty]$ be $\mathcal{S}$-measurable with $\int_X g \, d\mu < \infty$. Then for every $\varepsilon > 0$ there exists $E \in \mathcal{S}$ such that

$$\mu(E) < \infty \quad \text{and} \quad \int_{X \setminus E} g \, d\mu < \varepsilon.$$

(In words: up to ε loss in the integral, you can restrict g to a finite-measure set.)

(3.31) Dominated Convergence Theorem (DCT). Let $(X, \mathcal{S}, \mu)$ be a measure space. Suppose $f : X \to [-\infty, \infty]$ is $\mathcal{S}$-measurable and $f_1, f_2, \ldots$ are $\mathcal{S}$-measurable functions such that

$$f_k(x) \to f(x) \quad \text{for a.e. } x \in X.$$

If there exists an $\mathcal{S}$-measurable function $g : X \to [0, \infty]$ such that

$$\int_X g \, d\mu < \infty \quad \text{and} \quad |f_k(x)| \leq g(x) \ \text{ for all } k \in \mathbb{Z}^+ \text{ and a.e. } x \in X,$$

then

$$\lim_{k \to \infty} \int_X f_k \, d\mu \ = \int_X f \, d\mu.$$

Proof (sketch in two cases). Fix $E \in \mathcal{S}$. By triangle inequality and domination,

$$\begin{aligned}\Big|\int_X f_k\,d\mu - \int_X f\,d\mu\Big| &\le \Big|\int_{X\setminus E} f_k\,d\mu\Big| + \Big|\int_{X\setminus E} f\,d\mu\Big| + \Big|\int_E f_k\,d\mu - \int_E f\,d\mu\Big| \\ &\le \int_{X\setminus E} |f_k|\,d\mu + \int_{X\setminus E} |f|\,d\mu + \int_E |f_k - f|\,d\mu \\ &\le 2\int_{X\setminus E} g\,d\mu + \int_E |f_k - f|\,d\mu. \qquad (*)\end{aligned}$$

Case 1: $\mu(X) < \infty$. Given $\varepsilon > 0$, use (3.28) to choose $\delta > 0$ such that $\mu(B) < \delta \Rightarrow \int_B g\,d\mu < \varepsilon/4$. By Egorov, choose $E \in \mathcal{S}$ with $\mu(X \setminus E) < \delta$ such that $f_k \to f$ uniformly on E. Then the first term in $(*)$ is $\le 2(\varepsilon/4) = \varepsilon/2$. Also $\mu(E) \le \mu(X) < \infty$ and uniform convergence gives $\int_E |f_k - f|\,d\mu \le \mu(E)\sup_E |f_k - f| \to 0$, so for k large the second term is $< \varepsilon/2$. Hence the total is $< \varepsilon$.

Case 2: $\mu(X) = \infty$. Given $\varepsilon > 0$, use (3.29) to choose $E \in \mathcal{S}$ with $\mu(E) < \infty$ and $\int_{X\setminus E} g\,d\mu < \varepsilon/4$. Then $(*)$ implies

$$\Big|\int_X f_k\,d\mu - \int_X f\,d\mu\Big| \le \varepsilon/2 + \int_E |f_k - f|\,d\mu.$$

Now apply Case 1 on the finite-measure space $(E, \mathcal{S} \cap E, \mu)$ (still dominated by $g\mathbf{1}_E$) to conclude $\int_E |f_k - f|\,d\mu \to 0$, so the RHS is $< \varepsilon$ for k large. □

Remark. DCT generalizes earlier limit-interchange results: instead of assuming nonnegativity (MCT) or finite measure + uniform boundedness (BCT), it assumes pointwise domination by an integrable function.

Lebesgue vs. Riemann; notation

(3.34) Result (Riemann integrable $\Leftrightarrow$ continuous a.e.). Let $a < b$ and let $f : [a, b] \to \mathbb{R}$ be bounded. Then f is Riemann integrable if and only if

$$\lambda\big(\{x \in [a, b] : f \text{ is not continuous at } x\}\big) = 0,$$

and in that case the Riemann integral equals the Lebesgue integral:

$$\int_a^b f(x)\,dx \;=\; \int_{[a,b]} f\,d\lambda.$$

(3.39) Definition/Convention ("$\int_a^b$" denotes Lebesgue integral). Let $-\infty \le a < b \le \infty$ and let $f : (a, b) \to \mathbb{R}$ be Lebesgue measurable. Then:

- $\displaystyle\int_a^b f$ and $\displaystyle\int_a^b f(x)\,dx$ mean $\displaystyle\int_{(a,b)} f\,d\lambda$.
- $\displaystyle\int_b^a f := -\int_a^b f$ (so, e.g., $\int_a^b f = \int_a^c f + \int_c^b f$ when all terms are defined).

L^1 and approximation

(3.40) Definition ($\|f\|_1$, $L^1(\mu)$). Let $(X, \mathcal{S}, \mu)$ be a measure space and let $f : X \to [-\infty, \infty]$ be $\mathcal{S}$-measurable. The L^1-norm is

$$\|f\|_1 := \int_X |f| \, d\mu.$$

The Lebesgue space $L^1(\mu)$ is

$$L^1(\mu) := \{f : \ X \to \mathbb{R} \text{ measurable and } \|f\|_1 < \infty\}.$$

Example (3.42): ℓ^1. If μ is counting measure on $\mathbb{Z}^+$ and $x = (x_1, x_2, \dots)$ is a sequence of real numbers (viewed as a function on $\mathbb{Z}^+$), then

$$\|x\|_1 = \sum_{k=1}^{\infty} |x_k|.$$

In this case $L^1(\mu)$ is typically denoted ℓ^1, the set of sequences with $\sum_{k=1}^{\infty} |x_k| < \infty$.

(3.43) Result (Properties of the L^1-norm). If $f, g \in L^1(\mu)$ and $c \in \mathbb{R}$, then:

1. $\|f\|_1 \geq 0$.
2. $\|f\|_1 = 0 \iff f = 0$ a.e.
3. $\|cf\|_1 = |c| \, \|f\|_1$.
4. $\|f + g\|_1 \leq \|f\|_1 + \|g\|_1$.

(3.44) Result (Approximation by simple functions). If $f \in L^1(\mu)$, then for every $\varepsilon > 0$ there exists a simple function $g \in L^1(\mu)$ such that

$$\|f - g\|_1 < \varepsilon.$$

(In words: every L^1 function can be approximated in L^1-norm by functions taking only finitely many values.)

(3.45) Notation ($L^1(\mathbb{R})$). $L^1(\mathbb{R})$ denotes $L^1(\lambda)$ where λ is Lebesgue measure on $\mathbb{R}$ (either on Borel sets or Lebesgue measurable sets, as appropriate). When working in $L^1(\mathbb{R})$, $\|f\|_1$ means $\int_{\mathbb{R}} |f| \, d\lambda$.

(3.46) Definition (Step function). A *step function* is a function $g : \mathbb{R} \to \mathbb{R}$ of the form

$$g = \sum_{j=1}^{n} a_j \, \mathbf{1}_{I_j},$$

where $I_1, \dots, I_n$ are intervals in $\mathbb{R}$ and $a_1, \dots, a_n \in \mathbb{R}$. If the intervals are disjoint, then

$$\|g\|_1 = \sum_{j=1}^{n} |a_j| \, |I_j|,$$

where $|I_j|$ denotes the length of I_j. In particular, $g \in L^1(\mathbb{R})$ if all I_j are bounded. (Endpoints do not matter for L^1 since changing inclusion/exclusion changes the set only by measure 0.)

(3.47) Result (Approximation by step functions). If $f \in L^1(\mathbb{R})$, then for every $\varepsilon > 0$ there exists a step function $g \in L^1(\mathbb{R})$ such that

$$\|f - g\|_1 < \varepsilon.$$

Proof (outline). By (3.44), choose sets $A_1, \dots, A_n \subset \mathbb{R}$ (Borel/Lebesgue measurable) with $|A_k| < \infty$ and scalars $a_1, \dots, a_n$ such that

$$\Big\| f - \sum_{k=1}^{n} a_k \, \mathbf{1}_{A_k} \Big\|_1 < \varepsilon/2.$$

For each k, choose an open set $G_k \supset A_k$ with $|G_k \setminus A_k|$ arbitrarily small (outer regularity). Write G_k as a countable union of disjoint open intervals and choose a finite union E_k of bounded open intervals with

$$|G_k \setminus E_k| \text{ small, hence } \ \|\mathbf{1}_{A_k} - \mathbf{1}_{E_k}\|_1 = |A_k \triangle E_k| \text{ small.}$$

Choosing these errors so that $\sum_{k=1}^{n} |a_k| \, \|\mathbf{1}_{A_k} - \mathbf{1}_{E_k}\|_1 < \varepsilon/2$, we get

$$\begin{aligned} \Big\| f - \sum_{k=1}^{n} a_k \, \mathbf{1}_{E_k} \Big\|_1 &\le \Big\| f - \sum_{k=1}^{n} a_k \, \mathbf{1}_{A_k} \Big\|_1 + \Big\| \sum_{k=1}^{n} a_k (\mathbf{1}_{A_k} - \mathbf{1}_{E_k}) \Big\|_1 \\ &\le \varepsilon/2 + \sum_{k=1}^{n} |a_k| \, \|\mathbf{1}_{A_k} - \mathbf{1}_{E_k}\|_1 < \varepsilon. \end{aligned}$$

Since each E_k is a finite union of bounded intervals, $\sum_{k=1}^{n} a_k \, \mathbf{1}_{E_k}$ is a step function. □

(3.48) Result (Approximation by continuous functions). If $f \in L^1(\mathbb{R})$, then for every $\varepsilon > 0$ there exists a continuous function $g : \mathbb{R} \to \mathbb{R}$ such that

$$\|f - g\|_1 < \varepsilon \quad \text{and} \quad \{x \in \mathbb{R} : \ g(x) \neq 0\} \text{ is bounded.}$$

(So g can be taken continuous with bounded support.)

Proof (outline). Start with a step-function approximation $\sum_{k=1}^{n} a_k \, \mathbf{1}_{I_k}$ from (3.47), with I_k bounded intervals. For each bounded interval $I_k = [b_k, c_k]$ (or similar), choose a continuous function g_k supported in a slightly larger bounded interval such that $\|\mathbf{1}_{I_k} - g_k\|_1$ is as small as desired (e.g., smooth/continuous cutoff ramps near the endpoints). Then, using the triangle inequality,

$$\Big\| f - \sum_{k=1}^{n} a_k g_k \Big\|_1 \le \Big\| f - \sum_{k=1}^{n} a_k \, \mathbf{1}_{I_k} \Big\|_1 + \sum_{k=1}^{n} |a_k| \, \|\mathbf{1}_{I_k} - g_k\|_1,$$

and we choose the step approximation and the g_k so the RHS is $< \varepsilon$. Finally set $g := \sum_{k=1}^{n} a_k g_k$; it is continuous and nonzero only on a bounded set. □

Exercises 3B (continued)

2. Give an example of a sequence $f_1, f_2, \ldots$ of functions $f_k : \mathbb{Z}^+ \to [0,\infty)$ such that

$$\lim_{k\to\infty} f_k(m) = 0 \quad \text{for each } m \in \mathbb{Z}^+, \qquad \text{but} \qquad \lim_{k\to\infty} \int f_k\, d\mu = 1,$$

where μ is counting measure on $\mathbb{Z}^+$.

Example. Define

$$f_k(n) := \begin{cases} 1, & n = k, \\ 0, & n \neq k. \end{cases}$$

Fix m. Then $f_k(m) = 1$ only when $k = m$, and $f_k(m) = 0$ for all $k > m$, hence $\lim_{k\to\infty} f_k(m) = 0$. But

$$\int f_k\, d\mu = \sum_{n=1}^{\infty} f_k(n) = 1 \quad \text{for every } k,$$

so $\lim_{k\to\infty} \int f_k\, d\mu = 1$.

4(a). Suppose $(X, \mathcal{S}, \mu)$ is a measure space with $\mu(X) < \infty$. Suppose $f : X \to [0,\infty)$ is bounded and $\mathcal{S}$-measurable. Prove that

$$\int_X f\, d\mu = \inf\Big\{ \sum_{j=1}^{m} \mu(A_j) \sup_{x\in A_j} f(x) \;:\; A_1, \ldots, A_m \text{ is an } \mathcal{S}\text{-partition of } X \Big\}.$$

Proof. For a partition $P = \{A_1, \ldots, A_m\}$ of X, define the *upper sum*

$$U(f,P) := \sum_{j=1}^{m} \mu(A_j) \sup_{A_j} f, \qquad \text{where } \sup_{A_j} f := \sup\{f(x) : x \in A_j\}.$$

Also define the *upper simple function*

$$s_P(x) := \sum_{j=1}^{m} \big(\sup_{A_j} f\big) \mathbf{1}_{A_j}(x).$$

Then $s_P \geq f$ pointwise, hence

$$\int f\, d\mu \leq \int s_P\, d\mu = \sum_{j=1}^{m} \big(\sup_{A_j} f\big) \mu(A_j) = U(f,P).$$

Therefore $\displaystyle\int f\, d\mu \leq \inf_P U(f,P)$.

For the reverse inequality, let $M := \sup_X f < \infty$. For $n \in \mathbb{Z}^+$ partition $[0,M]$ into n subintervals and set

$$A_k^{(n)} := \Big\{ x \in X : \frac{(k-1)M}{n} < f(x) \leq \frac{kM}{n} \Big\}, \qquad k = 1, \ldots, n.$$

Then $\{A_1^{(n)}, \ldots, A_n^{(n)}\}$ is an $\mathcal{S}$-partition of X. Define

$$\psi_n(x) := \sum_{k=1}^{n} \frac{kM}{n} \mathbf{1}_{A_k^{(n)}}(x).$$

Then ψ_n is simple, $\psi_n \ge f$, and $0 \le \psi_n(x) - f(x) \le M/n$ for all x. Hence

$$0 \le \int \psi_n \, d\mu - \int f \, d\mu \le \int \frac{M}{n} \, d\mu = \frac{M}{n}\mu(X) \xrightarrow[n\to\infty]{} 0.$$

Moreover, on each $A_k^{(n)}$ we have $\sup_{A_k^{(n)}} f \le kM/n$, so

$$U(f, P_n) = \sum_{k=1}^{n} \mu\big(A_k^{(n)}\big) \sup_{A_k^{(n)}} f \le \sum_{k=1}^{n} \mu\big(A_k^{(n)}\big) \frac{kM}{n} = \int \psi_n \, d\mu.$$

Thus for any $\varepsilon > 0$, for n large we have

$$U(f, P_n) \le \int \psi_n \, d\mu \le \int f \, d\mu + \varepsilon,$$

so $\inf_P U(f, P) \le \int f \, d\mu$. Combine with the first inequality to get equality.

10(a). Suppose $(X, \mathcal{S}, \mu)$ is a measure space with $\mu(X) < \infty$, $0 < p < r$, and $f : X \to [0, \infty)$ is $\mathcal{S}$-measurable. Prove that if

$$\int_X f^r \, d\mu < \infty,$$

then $\int_X f^p \, d\mu < \infty$.

Proof. Split X into

$$A := \{x : f(x) < 1\}, \qquad B := \{x : f(x) \ge 1\}.$$

On A, $f^p \le 1$, and on B, since $f \ge 1$ and $p < r$, we have $f^p \le f^r$. Hence

$$\int_X f^p \, d\mu = \int_A f^p \, d\mu + \int_B f^p \, d\mu \le \mu(A) + \int_B f^r \, d\mu \le \mu(X) + \int_X f^r \, d\mu < \infty.$$

10(b). Give an example showing (a) can fail if $\mu(X) = \infty$.

Example. Let $X = \mathbb{Z}^+$ with $\mathcal{S} = \mathcal{P}(\mathbb{Z}^+)$ and $\mu =$ counting measure, and choose β such that

$$\beta r > 1 \quad \text{but} \quad \beta p \le 1 \qquad (\text{e.g. } \beta = \tfrac{1}{2}(\tfrac{1}{r} + \tfrac{1}{p})).$$

Define $f(n) = n^{-\beta}$. Then

$$\int f^r \, d\mu = \sum_{n=1}^{\infty} n^{-\beta r} < \infty \quad (\beta r > 1), \qquad \text{but} \qquad \int f^p \, d\mu = \sum_{n=1}^{\infty} n^{-\beta p} = \infty \quad (\beta p \le 1).$$

Chapter 4: Differentiation

Section 4A: Hardy–Littlewood Maximal Function

Result (Markov's inequality). If $(X, \mathcal{S}, \mu)$ is a measure space and $h \in L^1(\mu)$, then for every $c > 0$,

$$\mu\big(\{x \in X : |h(x)| \ge c\}\big) \le \frac{1}{c} \, \|h\|_1.$$

Proof. On the set $\{|h| \ge c\}$ we have $c \le |h|$, so

$$c\, \mu(\{|h| \ge c\}) = \int_{\{|h| \ge c\}} c \, d\mu \le \int_{\{|h| \ge c\}} |h| \, d\mu \le \int_X |h| \, d\mu = \|h\|_1.$$

Definition ($3I$). If $I \subset \mathbb{R}$ is a bounded, nonempty open interval, then $3I$ denotes the open interval with the same center as I and three times the length of I.

Lemma (Vitali covering lemma, finite version). If $I_1, \ldots, I_n$ is a finite list of bounded, nonempty open intervals in $\mathbb{R}$, then there exists a *disjoint* sublist $I_{k_1}, \ldots, I_{k_m}$ such that

$$I_1 \cup \cdots \cup I_n \subset (3I_{k_1}) \cup \cdots \cup (3I_{k_m}).$$

Definition (Hardy–Littlewood maximal function, centered). If $h : \mathbb{R} \to \mathbb{R}$ is Lebesgue measurable, define $h^* : \mathbb{R} \to [0, \infty]$ by

$$h^*(b) := \sup_{t>0} \frac{1}{2t} \int_{b-t}^{b+t} |h(x)| \, dx.$$

Equivalently, $h^*(b)$ is the supremum over all bounded intervals centered at b of the average of $|h|$ on the interval.

Result (Hardy–Littlewood maximal inequality). If $h \in L^1(\mathbb{R})$, then for every $c > 0$,

$$\big|\{b \in \mathbb{R} : h^*(b) > c\}\big| \leq \frac{3}{c} \, \|h\|_1,$$

where $|\cdot|$ denotes Lebesgue measure.

Proof (outline). Let F be a closed, bounded subset of $\{h^* > c\}$. For each $b \in F$ choose $t_b > 0$ so that

$$\frac{1}{2t_b} \int_{b-t_b}^{b+t_b} |h| > c.$$

Then $\{(b - t_b, b + t_b)\}_{b \in F}$ is an open cover of F, hence has a finite subcover $I_1, \ldots, I_n$ by Heine–Borel. Apply Vitali's lemma to obtain disjoint intervals $I_{k_1}, \ldots, I_{k_m}$ with

$$F \subset I_1 \cup \cdots \cup I_n \subset (3I_{k_1}) \cup \cdots \cup (3I_{k_m}).$$

Therefore,

$$|F| \leq \sum_{j=1}^{m} |3I_{k_j}| = 3 \sum_{j=1}^{m} |I_{k_j}|.$$

But each chosen interval satisfies $\int_{I_{k_j}} |h| > c\,|I_{k_j}|$, hence $|I_{k_j}| \leq \frac{1}{c} \int_{I_{k_j}} |h|$. Summing and using disjointness,

$$|F| \leq \frac{3}{c} \sum_{j=1}^{m} \int_{I_{k_j}} |h| \leq \frac{3}{c} \int_{\mathbb{R}} |h| = \frac{3}{c} \|h\|_1.$$

Taking the supremum over closed bounded $F \subset \{h^* > c\}$ yields the claim for the whole set.

Exercises 4A (notes).

2. (Chebyshev / variance form). Assume $(X, \mathcal{S}, \mu)$ is a measure space with $\mu(X) = 1$ and $h \in L^2(\mu)$. Show that for all $c > 0$,

$$\mu\Big(\big\{x : |h(x) - \textstyle\int h \, d\mu| \geq c\big\}\Big) \leq \tfrac{1}{c^2}\Big(\|h\|_2^2 - \big(\textstyle\int h \, d\mu\big)^2\Big).$$

Proof. Let $m := \int h\,d\mu$ and $A := \{x : |h-m| \ge c\}$. On A, $(h-m)^2 \ge c^2$, hence

$$c^2\mu(A) \le \int_A (h-m)^2\,d\mu \le \int_X (h-m)^2\,d\mu.$$

Expand:

$$\int (h-m)^2 = \int h^2 - 2m\int h + m^2 \int 1 = \int h^2 - 2m^2 + m^2\mu(X) = \int h^2 - m^2,$$

since $\mu(X) = 1$. Divide by c^2.

9. (Openness of the superlevel set; non-centered variant). Let $h : \mathbb{R} \to \mathbb{R}$ be Lebesgue measurable and define the *non-centered* maximal function

$$h^*(x) := \sup_{I \ni x} \frac{1}{|I|} \int_I |h|,$$

where the supremum is over bounded intervals $I \subset \mathbb{R}$ containing x. Then for every $c \in \mathbb{R}$,

$$\{b \in \mathbb{R} : h^*(b) > c\} \text{ is open.}$$

Proof. Let $E_c := \{x : h^*(x) > c\}$ and take $b \in E_c$. By definition of supremum, there exists an interval I with $b \in I$ and $\frac{1}{|I|}\int_I |h| > c$. For any $x \in I$, the same interval I contains x, so $h^*(x) \ge \frac{1}{|I|}\int_I |h| > c$. Hence $I \subset E_c$, so every $b \in E_c$ has a neighborhood contained in E_c and E_c is open.

Section 4B: Derivatives of Integrals

(V1) Lebesgue Differentiation Theorem. If $f \in L^1(\mathbb{R})$, then

$$\lim_{t \downarrow 0} \frac{1}{2t} \int_{b-t}^{b+t} |f(x) - f(b)|\,dx = 0 \quad \text{for a.e. } b \in \mathbb{R}.$$

Note (continuous-point special case). For $t > 0$, by the integral bound (cf. $\int_E \le \mu(E)\sup_E$),

$$\frac{1}{2t} \int_{b-t}^{b+t} |f(x) - f(b)|\,dx \le \sup\{|f(x) - f(b)| : |x - b| \le t\}.$$

If f is continuous at b, the RHS $\to 0$ as $t \to 0$, proving the theorem at such b.

Definition (Derivative). Let $g : I \to \mathbb{R}$ be defined on an open interval $I \subset \mathbb{R}$ and $b \in I$. The derivative of g at b is

$$g'(b) := \lim_{t \to 0} \frac{g(b+t) - g(b)}{t},$$

if the limit exists; in that case g is differentiable at b.

Fundamental Theorem of Calculus (continuous point). Suppose $f \in L^1(\mathbb{R})$ and define $g : \mathbb{R} \to \mathbb{R}$ by

$$g(x) := \int_{-\infty}^{x} f.$$

If f is continuous at b, then $g'(b) = f(b)$.

Proof. For $t \neq 0$,

$$\frac{g(b+t)-g(b)}{t} - f(b) = \frac{1}{t}\int_b^{b+t} \big(f(x)-f(b)\big)\,dx,$$

so

$$\left|\frac{g(b+t)-g(b)}{t} - f(b)\right| \le \sup\{|f(x)-f(b)| : |x-b| \le |t|\} \xrightarrow[t\to 0]{} 0.$$

(V2) Lebesgue Differentiation Theorem (derivative form). If $f \in L^1(\mathbb{R})$ and $g(x) = \int_{-\infty}^x f$, then

$$g'(b) = f(b) \quad \text{for a.e. } b \in \mathbb{R}.$$

Result ("No set is exactly half of every initial interval"). There does not exist a Lebesgue measurable set $E \subset [0,1]$ such that

$$|E \cap [0,b]| = \frac{b}{2} \quad \text{for all } b \in [0,1].$$

Result (L^1 function equals its local average a.e.). If $f \in L^1(\mathbb{R})$, then

$$f(b) = \lim_{t\downarrow 0} \frac{1}{2t}\int_{b-t}^{b+t} f(x)\,dx \quad \text{for a.e. } b \in \mathbb{R}.$$

(In words: f agrees a.e. with the limit of its symmetric local averages.)

Definition (Density). For $E \subset \mathbb{R}$ and $b \in \mathbb{R}$, the *density of E at b* is

$$\lim_{t\downarrow 0} \frac{|E \cap (b-t,b+t)|}{2t},$$

if the limit exists (otherwise the density at b is undefined).
Example (density of $[0,1]$*).*

$$\operatorname{dens}_{[0,1]}(b) = \begin{cases} 1, & b \in (0,1), \\ \frac{1}{2}, & b = 0 \text{ or } b = 1, \\ 0, & b \notin [0,1]. \end{cases}$$

Lebesgue Density Theorem. If $E \subset \mathbb{R}$ is Lebesgue measurable, then the density of E equals 1 at a.e. point of E and equals 0 at a.e. point of $\mathbb{R} \setminus E$.

Result ("Bad" Borel set). There exists a Borel set $E \subset \mathbb{R}$ such that for every nonempty bounded open interval I,

$$0 < |E \cap I| < |I|.$$

(*Remark.* This does not contradict the density theorem: density statements are a.e., while the condition above is pointwise for *every* interval.)

Exercises 4B (notes). For $f \in L^1(\mathbb{R})$ and an interval $I \subset \mathbb{R}$ with $0 < |I| < \infty$, let

$$f_I := \frac{1}{|I|}\int_I f \qquad \text{(the average of } f \text{ on } I).$$

1. If $f \in L^1(\mathbb{R})$, prove that

$$\lim_{t\downarrow 0} \frac{1}{2t}\int_{b-t}^{b+t} \left|f(x) - f_{[b-t,b+t]}\right| dx = 0 \quad \text{for a.e. } b \in \mathbb{R}.$$

Proof. For any interval I with $0 < |I| < \infty$ and any constant c,

$$\begin{aligned}\int_I |f - f_I| \le \int_I \big(|f-c| + |f_I - c|\big) \le \int_I |f-c| + |I| \cdot \frac{1}{|I|}\int_I |f-c| \\ = 2\int_I |f-c|.\end{aligned}$$

Take $I = [b-t, b+t]$ and $c = f(b)$ to get

$$\frac{1}{2t}\int_{b-t}^{b+t} |f - f_I| \le 2 \cdot \frac{1}{2t}\int_{b-t}^{b+t} |f - f(b)|.$$

The RHS $\to 0$ for a.e. b by (V1), hence so does the LHS.

2. If $f \in L^1(\mathbb{R})$, prove that for a.e. $b \in \mathbb{R}$,

$$\lim_{t\downarrow 0} \sup\left\{\frac{1}{|I|}\int_I |f - f_I| : \ I \text{ is an interval of length } t \text{ containing } b\right\} = 0.$$

Proof. Fix b in the full-measure set where (V1) holds. For any interval I of length t containing b, using the inequality from (1) with $c = f(b)$,

$$\frac{1}{|I|}\int_I |f - f_I| \le \frac{2}{|I|}\int_I |f - f(b)| \le \frac{2}{t}\int_{b-t}^{b+t} |f - f(b)| = 4 \cdot \frac{1}{2t}\int_{b-t}^{b+t} |f - f(b)|.$$

The RHS $\to 0$ as $t \downarrow 0$, so the supremum (bounded by the same RHS) also tends to 0.

3. Suppose $f : \mathbb{R} \to \mathbb{R}$ is Lebesgue measurable and $f^2 \in L^1(\mathbb{R})$. Prove that

$$\lim_{t\downarrow 0} \frac{1}{2t}\int_{b-t}^{b+t} \left|f(x) - f(b)\right|^2 dx = 0 \quad \text{for a.e. } b \in \mathbb{R}.$$

Proof (outline). Use the elementary inequality $u^2 + 1 \ge 2|u|$, i.e.

$$|u| \le \frac{u^2+1}{2}.$$

On any bounded interval I this gives

$$\int_I |f| \le \frac{1}{2}\int_I f^2 + \frac{1}{2}|I| < \infty,$$

so $f \in L^1_{\text{loc}}(\mathbb{R})$ and the Lebesgue differentiation theorem applies to both f and f^2. Hence there is a full-measure set E such that for every $b \in E$,

$$\lim_{t\downarrow 0} \frac{1}{2t}\int_{b-t}^{b+t} f(x)\, dx = f(b), \qquad \lim_{t\downarrow 0} \frac{1}{2t}\int_{b-t}^{b+t} f(x)^2\, dx = f(b)^2.$$

Fix $b \in E$. Then

$$\begin{aligned}\frac{1}{2t}\int_{b-t}^{b+t} (f(x) - f(b))^2\, dx &= \frac{1}{2t}\int_{b-t}^{b+t} \big(f(x)^2 - 2f(b)f(x) + f(b)^2\big)\, dx \\ &= \underbrace{\frac{1}{2t}\int_{b-t}^{b+t} f(x)^2\, dx}_{A_t} - 2f(b)\underbrace{\frac{1}{2t}\int_{b-t}^{b+t} f(x)\, dx}_{B_t} + f(b)^2.\end{aligned}$$

As $t \downarrow 0$, $A_t \to f(b)^2$ and $B_t \to f(b)$, so the expression tends to $f(b)^2 - 2f(b)\cdot f(b) + f(b)^2 = 0$.

Exercises 4B (continued)

4. (LDT under local integrability). Show that the Lebesgue Differentiation Theorem still holds if the hypothesis $f \in L^1(\mathbb{R})$ is weakened to

$$f \in L^1_{\text{loc}}(\mathbb{R}) \quad (\text{equivalently: } \int_{b-t}^{b+t} |f(x)|\,dx < \infty \text{ for all } b \in \mathbb{R},\ t > 0).$$

Proof. For $n \in \mathbb{Z}^+$ define $f_n := f\,\mathbf{1}_{[-n,n]}$. Then $f_n \in L^1(\mathbb{R})$, so by the (global) Lebesgue Differentiation Theorem,

$$\lim_{t\downarrow 0} \frac{1}{2t} \int_{b-t}^{b+t} \big|f_n(x) - f_n(b)\big|\,dx = 0 \quad \text{for a.e. } b \in \mathbb{R}.$$

Let E_n be the full-measure set where this holds, and set $E := \bigcap_{n=1}^{\infty} E_n$ (still full measure). Fix $b \in E$ and choose $n > |b| + 1$. Then for all sufficiently small $t > 0$ we have $[b-t, b+t] \subset (-n, n)$, so $f_n(x) = f(x)$ for $x \in [b-t, b+t]$ and $f_n(b) = f(b)$. Hence for such t,

$$\frac{1}{2t} \int_{b-t}^{b+t} \big|f(x) - f(b)\big|\,dx = \frac{1}{2t} \int_{b-t}^{b+t} \big|f_n(x) - f_n(b)\big|\,dx \xrightarrow[t\downarrow 0]{} 0.$$

Thus the LDT conclusion holds for f at every $b \in E$, i.e. for a.e. b.

5. Suppose $f : \mathbb{R} \to \mathbb{R}$ is Lebesgue measurable. Prove that

$$|f(b)| \le f^*(b) \quad \text{for a.e. } b \in \mathbb{R},$$

where $f^*(b) = \sup_{t>0} \frac{1}{2t} \int_{b-t}^{b+t} |f(x)|\,dx$ is the (centered) Hardy–Littlewood maximal function.

Proof. Apply the Lebesgue Differentiation Theorem to $|f|$ (note $|f| \in L^1_{\text{loc}}$ whenever the averages are finite; otherwise $f^*(b) = \infty$ and the inequality is trivial). For a.e. b,

$$|f(b)| = \lim_{t\downarrow 0} \frac{1}{2t} \int_{b-t}^{b+t} |f(x)|\,dx.$$

Since $f^*(b)$ is the supremum over all $t > 0$ of these averages, we have $f^*(b) \ge$ (the limit of the averages), hence $|f(b)| \le f^*(b)$ for a.e. b.

6. Prove that if $h \in L^1(\mathbb{R})$ and

$$\int_{-\infty}^{x} h(t)\,dt = 0 \quad \text{for all } x \in \mathbb{R},$$

then $h(x) = 0$ for a.e. $x \in \mathbb{R}$.

Proof. Fix $b \in \mathbb{R}$ and $t > 0$. Then

$$\frac{1}{2t} \int_{b-t}^{b+t} h(x)\,dx = \frac{1}{2t} \Big(\int_{-\infty}^{b+t} h - \int_{-\infty}^{b-t} h \Big) = 0$$

by the assumption. By the Lebesgue Differentiation Theorem applied to $h \in L^1(\mathbb{R})$,

$$h(b) = \lim_{t\downarrow 0} \frac{1}{2t} \int_{b-t}^{b+t} h(x)\,dx = 0 \quad \text{for a.e. } b \in \mathbb{R}.$$

Chapter 5: Product Measure

Section 5A: Products of Measure Spaces

(5.1) Definition (Rectangle). Let X, Y be sets. A *rectangle* in $X \times Y$ is a set of the form $A \times B$ where $A \subset X$ and $B \subset Y$. (Note: A and B need not be intervals.)

(5.2) Definition (Product σ-algebra $\mathcal{S} \otimes \mathcal{T}$). Suppose $(X, \mathcal{S})$ and $(Y, \mathcal{T})$ are measurable spaces. Then:

(i) $\mathcal{S} \otimes \mathcal{T}$ is the *smallest* σ-algebra on $X \times Y$ that contains all measurable rectangles:

$$\{A \times B : \ A \in \mathcal{S}, \ B \in \mathcal{T}\}.$$

(ii) A *measurable rectangle* (with respect to $\mathcal{S} \otimes \mathcal{T}$) is any set $A \times B$ with $A \in \mathcal{S}$, $B \in \mathcal{T}$.

(5.3) Definition (Cross sections of sets). Let X, Y be sets and $E \subset X \times Y$. For $a \in X$ and $b \in Y$ define the cross sections

$$[E]_a := \{y \in Y : \ (a, y) \in E\}, \qquad [E]^b := \{x \in X : \ (x, b) \in E\}.$$

Example. If $A \subset X$ and $B \subset Y$, then for $a \in X$, $b \in Y$,

$$[A \times B]_a = \begin{cases} B, & a \in A, \\ \varnothing, & a \notin A, \end{cases} \qquad [A \times B]^b = \begin{cases} A, & b \in B, \\ \varnothing, & b \notin B. \end{cases}$$

(5.6) Result (Cross sections preserve measurability for sets). If $\mathcal{S}$ is a σ-algebra on X, $\mathcal{T}$ is a σ-algebra on Y, and $E \in \mathcal{S} \otimes \mathcal{T}$, then

$$[E]_a \in \mathcal{T} \ \ \forall a \in X, \qquad \text{and} \qquad [E]^b \in \mathcal{S} \ \ \forall b \in Y.$$

(5.7) Definition (Cross sections of functions). Let X, Y be sets and $f : X \times Y \to \mathbb{R}$ a function. For $a \in X$ and $b \in Y$, define the cross-section functions

$$[f]_a : Y \to \mathbb{R}, \quad [f]_a(y) := f(a, y), \qquad [f]^b : X \to \mathbb{R}, \quad [f]^b(x) := f(x, b).$$

Examples.

(i) If $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $f(x, y) = 5x^2 + y^3$, then

$$[f]_2(y) = f(2, y) = 20 + y^3, \qquad [f]^3(x) = f(x, 3) = 5x^2 + 27.$$

(ii) If $A \subset X$, $B \subset Y$, and $f = \mathbf{1}_{A \times B}$, then

$$[f]_a(y) = \mathbf{1}_{A \times B}(a, y) = \mathbf{1}_A(a)\mathbf{1}_B(y) = \mathbf{1}_A(a)\,\mathbf{1}_B(y),$$

so $[f]_a = \mathbf{1}_A(a)\,\mathbf{1}_B$, and similarly $[f]^b = \mathbf{1}_B(b)\,\mathbf{1}_A$.

(5.8) Result (Cross sections preserve measurability for functions). If $\mathcal{S}$ is a σ-algebra on X, $\mathcal{T}$ is a σ-algebra on Y, and $f : X \times Y \to \mathbb{R}$ is $(\mathcal{S} \otimes \mathcal{T})$-measurable, then:

(i) For every $a \in X$, $[f]_a$ is $\mathcal{T}$-measurable on Y.

(ii) For every $b \in Y$, $[f]^b$ is $\mathcal{S}$-measurable on X.

Section 5A (continued)

Standard 2-step proof technique (for properties of σ-algebras). To show that *every* set in $\sigma(\mathcal{G})$ has some property P:

1. Show every set in the generating class $\mathcal{G}$ has property P.
2. Show the collection $\mathcal{C} := \{E \in \sigma(\mathcal{G}) : E \text{ has } \mathsf{P}\}$ is a σ-algebra.

Then $\mathcal{G} \subseteq \mathcal{C}$ and $\mathcal{C}$ is a σ-algebra, so $\sigma(\mathcal{G}) \subseteq \mathcal{C}$.

(5.10) Definition (Algebra). Let W be a set and let $\mathcal{A} \subseteq \mathcal{P}(W)$. We say $\mathcal{A}$ is an *algebra* on W if:

(i) $\varnothing \in \mathcal{A}$,

(ii) if $E \in \mathcal{A}$, then $W \setminus E \in \mathcal{A}$,

(iii) if $E, F \in \mathcal{A}$, then $E \cup F \in \mathcal{A}$.

Equivalently: an algebra is closed under complements (relative to W) and *finite* unions. A σ-algebra is closed under complements and *countable* unions.

Examples.

- The collection of all finite unions of intervals in $\mathbb{R}$ is an algebra (but not a σ-algebra).
- The collection of all countable unions of intervals in $\mathbb{R}$ is closed under countable unions, but (in general) not under complements, so it need not be a σ-algebra.

(5.13) Result (Finite unions of measurable rectangles form an algebra). Suppose $(X, \mathcal{S})$ and $(Y, \mathcal{T})$ are measurable spaces.

(a) The set of all finite unions of measurable rectangles in $X \times Y$

$$\left\{ \bigcup_{j=1}^{n} (A_j \times B_j) : \; n \in \mathbb{N}, \; A_j \in \mathcal{S}, \; B_j \in \mathcal{T} \right\}$$

is an algebra on $X \times Y$.

(b) Every finite union of measurable rectangles can be written as a *finite union of disjoint* measurable rectangles.

(5.15) Definition (Monotone class). Let W be a set and $\mathcal{M} \subseteq \mathcal{P}(W)$. We say $\mathcal{M}$ is a *monotone class* on W if:

(i) If $E_1 \subseteq E_2 \subseteq \cdots$ with each $E_n \in \mathcal{M}$, then $\bigcup_{n=1}^{\infty} E_n \in \mathcal{M}$.

(ii) If $E_1 \supseteq E_2 \supseteq \cdots$ with each $E_n \in \mathcal{M}$, then $\bigcap_{n=1}^{\infty} E_n \in \mathcal{M}$.

Notes:

- Every σ-algebra is a monotone class.

- Some monotone classes are *not* closed under finite unions or complements.
- If $\mathcal{A} \subseteq \mathcal{P}(W)$, then the intersection of all monotone classes containing $\mathcal{A}$ is the *smallest* monotone class containing $\mathcal{A}$.

(5.17) Result (Monotone Class Theorem). If $\mathcal{A}$ is an algebra on W, then the smallest σ-algebra containing $\mathcal{A}$ equals the smallest monotone class containing $\mathcal{A}$. (In particular, this is useful for proving properties for all sets in $\sigma(\mathcal{A})$.)

(5.18) Definition (Finite and σ-finite measures). Let $(X, \mathcal{S}, \mu)$ be a measure space.

- μ is *finite* if $\mu(X) < \infty$.
- μ is *σ-finite* if there exist $X_1, X_2, \dots \in \mathcal{S}$ such that

$$X = \bigcup_{k=1}^{\infty} X_k \quad \text{and} \quad \mu(X_k) < \infty \ \ \forall k.$$

Examples:

- Lebesgue measure on $[0, 1]$ is finite.
- Lebesgue measure on $\mathbb{R}$ is not finite, but is σ-finite.
- Counting measure on $\mathbb{R}$ is not σ-finite (a countable union of finite sets is countable).

(5.20) Result (Measure of cross-sections is measurable). Suppose $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{T}, \nu)$ are σ-finite measure spaces. If $E \in \mathcal{S} \otimes \mathcal{T}$, then

(a) $x \mapsto \nu([E]_x)$ is $\mathcal{S}$-measurable on X,

(b) $y \mapsto \mu([E]^y)$ is $\mathcal{T}$-measurable on Y.

(5.21) Definition (Integration notation). If $(X, \mathcal{S}, \mu)$ is a measure space and $g : X \to [-\infty, \infty]$, then

$$\int g(x)\, d\mu(x)$$

means $\int g\, d\mu$; the $d\mu(x)$ notation emphasizes that variables other than x are treated as constants. *Example (Lebesgue λ).* For fixed x,

$$\int_{[0,4]} (x^2 + y)\, d\lambda(y) = \int_0^4 x^2\, dy + \int_0^4 y\, dy = 4x^2 + 8.$$

For fixed y,

$$\int_{[0,4]} (x^2 + y)\, d\lambda(x) = \int_0^4 x^2\, dx + \int_0^4 y\, dx = \frac{64}{3} + 4y.$$

(5.23) Definition (Iterated integrals). Suppose $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{T}, \nu)$ are measure spaces and $f : X \times Y \to \mathbb{R}$. Then

$$\int_X \left(\int_Y f(x, y)\, d\nu(y) \right) d\mu(x)$$

is often written as

$$\int_X \int_Y f(x,y)\, d\nu(y)\, d\mu(x).$$

In words: first fix x and integrate $y \mapsto f(x,y)$ against ν, producing a function of x; then integrate that function against μ.

Example. Over $[0,4] \times [0,4]$ (Lebesgue),

$$\int_0^4 \int_0^4 (x^2 + y)\, dy\, dx = \int_0^4 (4x^2 + 8)\, dx = \frac{352}{3}.$$

Similarly,

$$\int_0^4 \int_0^4 (x^2 + y)\, dx\, dy = \int_0^4 \left(\frac{64}{3} + 4y \right) dy = \frac{352}{3}.$$

(5.25) Definition (Product of two measures). Suppose $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{T}, \nu)$ are σ-finite measure spaces. For $E \in \mathcal{S} \otimes \mathcal{T}$, define

$$(\mu \times \nu)(E) := \int_X \int_Y \mathbf{1}_E(x,y)\, d\nu(y)\, d\mu(x).$$

Example (measurable rectangle). If $A \in \mathcal{S}$ and $B \in \mathcal{T}$, then

$$(\mu \times \nu)(A \times B) = \int_X \int_Y \mathbf{1}_A(x) \mathbf{1}_B(y)\, d\nu(y)\, d\mu(x) = \mu(A)\nu(B).$$

(5.27) Result (Product of two measures is a measure). If $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{T}, \nu)$ are σ-finite measure spaces, then $\mu \times \nu$ is a measure on $(X \times Y, \mathcal{S} \otimes \mathcal{T})$.

Proof sketch. Trivially $(\mu \times \nu)(\varnothing) = 0$. If $E_1, E_2, \ldots$ are disjoint in $\mathcal{S} \otimes \mathcal{T}$, then $\mathbf{1}_{\cup_k E_k} = \sum_{k=1}^\infty \mathbf{1}_{E_k}$. By monotone convergence,

$$(\mu \times \nu)\Big(\bigcup_{k=1}^\infty E_k \Big) = \int_X \int_Y \sum_{k=1}^\infty \mathbf{1}_{E_k}\, d\nu\, d\mu = \sum_{k=1}^\infty \int_X \int_Y \mathbf{1}_{E_k}\, d\nu\, d\mu = \sum_{k=1}^\infty (\mu \times \nu)(E_k).$$

Hence $\mu \times \nu$ is countably additive.

Exercise 5A.2. Suppose $(X, \mathcal{S})$ and $(Y, \mathcal{T})$ are measurable spaces. Prove that if $A \neq \varnothing$, $A \subseteq X$ and $B \neq \varnothing$, $B \subseteq Y$ satisfy $A \times B \in \mathcal{S} \otimes \mathcal{T}$, then $A \in \mathcal{S}$ and $B \in \mathcal{T}$.

Proof. For $E \subseteq X \times Y$ and $x \in X$ write $E_x := [E]_x = \{y \in Y : (x,y) \in E\}$. Fix $x \in X$ and define

$$\mathcal{C}_x := \{E \subseteq X \times Y : \ E_x \in \mathcal{T}\}.$$

Then $\mathcal{C}_x$ is a σ-algebra on $X \times Y$ since

$$(E^c)_x = (E_x)^c, \qquad \left(\bigcup_{n=1}^\infty E_n \right)_x = \bigcup_{n=1}^\infty (E_n)_x.$$

Moreover, for any measurable rectangle $S \times T$ with $S \in \mathcal{S}$, $T \in \mathcal{T}$,

$$(S \times T)_x = \begin{cases} T, & x \in S, \\ \varnothing, & x \notin S, \end{cases} \in \mathcal{T},$$

so every measurable rectangle lies in $\mathcal{C}_x$. Hence $\mathcal{S} \otimes \mathcal{T} \subseteq \mathcal{C}_x$, and therefore:

$$E \in \mathcal{S} \otimes \mathcal{T} \implies E_x \in \mathcal{T} \quad \forall x \in X.$$

Applying this to $E = A \times B$, pick $x_0 \in A$ (since $A \neq \varnothing$). Then

$$(A \times B)_{x_0} = B \in \mathcal{T}.$$

Similarly, using $E^y := [E]^y$ and the symmetric argument (fix y and consider $\{E : E^y \in \mathcal{S}\}$), we get

$$E \in \mathcal{S} \otimes \mathcal{T} \implies E^y \in \mathcal{S} \quad \forall y \in Y.$$

Pick $y_0 \in B$. Then $(A \times B)^{y_0} = A \in \mathcal{S}$. $\square$

Chapter 5, Section 5B: Iterated Integrals

(5.28) Result (Tonelli's Theorem). Suppose $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{T}, \nu)$ are σ-finite measure spaces and $f : X \times Y \to [0, \infty]$ is $(\mathcal{S} \otimes \mathcal{T})$-measurable. Then:

(a) The function $x \mapsto \displaystyle\int_Y f(x, y)\, d\nu(y)$ is $\mathcal{S}$-measurable on X.

(b) The function $y \mapsto \displaystyle\int_X f(x, y)\, d\mu(x)$ is $\mathcal{T}$-measurable on Y.

(c) The integrals satisfy

$$\int_{X \times Y} f\, d(\mu \times \nu) = \int_X \left(\int_Y f(x, y)\, d\nu(y) \right) d\mu(x) = \int_Y \left(\int_X f(x, y)\, d\mu(x) \right) d\nu(y),$$

allowing the order of integration to be switched.

Interpretation. If $f \geq 0$ and the measures are σ-finite, then iterated integrals always exist (possibly $+\infty$) and can be interchanged.

Example (Tonelli can fail without σ-finiteness). Let $\mathcal{B}$ be the Borel σ-algebra on $[0, 1]$, let λ be Lebesgue measure on $([0, 1], \mathcal{B})$, and let μ be counting measure on $([0, 1], \mathcal{B})$ (not σ-finite). Let $D = \{(x, x) : x \in [0, 1]\}$ be the diagonal and $f = \mathbf{1}_D$. Then

$$\int_{[0,1]} \left(\int_{[0,1]} \mathbf{1}_D(x, y)\, d\mu(y) \right) d\lambda(x) = \int_{[0,1]} 1\, d\lambda(x) = 1,$$

but for each fixed y, the set $\{x : (x, y) \in D\} = \{y\}$ has λ-measure 0, so

$$\int_{[0,1]} \left(\int_{[0,1]} \mathbf{1}_D(x, y)\, d\lambda(x) \right) d\mu(y) = \int_{[0,1]} 0\, d\mu(y) = 0.$$

(5.30) Result (Double sums of nonnegative numbers). If $\{x_{j,k} : j, k \in \mathbb{Z}^+\}$ is a doubly indexed collection of nonnegative numbers, then

$$\sum_{j=1}^{\infty} \sum_{k=1}^{\infty} x_{j,k} = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} x_{j,k} \quad \text{(possibly } +\infty\text{)}.$$

(Think: Tonelli for counting measure.)

(5.32) Result (Fubini's Theorem). Suppose $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{T}, \nu)$ are σ-finite measure spaces and $f : X \times Y \to [-\infty, \infty]$ is $(\mathcal{S} \otimes \mathcal{T})$-measurable with

$$\int_{X \times Y} |f| \, d(\mu \times \nu) < \infty.$$

Then:

(a) For a.e. $x \in X$, $\displaystyle\int_Y |f(x, y)| \, d\nu(y) < \infty$.

(b) For a.e. $y \in Y$, $\displaystyle\int_X |f(x, y)| \, d\mu(x) < \infty$.

(c) The functions $x \mapsto \int_Y f(x, y) \, d\nu(y)$ and $y \mapsto \int_X f(x, y) \, d\mu(x)$ are measurable (a.e. defined), and

$$\int_{X \times Y} f \, d(\mu \times \nu) = \int_X \left(\int_Y f(x, y) \, d\nu(y) \right) d\mu(x) = \int_Y \left(\int_X f(x, y) \, d\mu(x) \right) d\nu(y).$$

Rule of thumb. Tonelli: $f \geq 0$. Fubini: $\int |f| < \infty$. In practice, prove $\int |f| < \infty$ by applying Tonelli to $|f|$.

(5.34) Definition (Region under a graph). Let X be a set and $f : X \to [0, \infty]$ be a function. The *region under the graph* of f is

$$U_f := \{(x, t) \in X \times [0, \infty) : \ 0 \leq t < f(x)\}.$$

(5.35) Result (Area under the graph equals the integral). Suppose $(X, \mathcal{S}, \mu)$ is a σ-finite measure space and $f : X \to [0, \infty]$ is $\mathcal{S}$-measurable. Let $\mathcal{B}$ be the Borel σ-algebra on $[0, \infty)$ and let λ be Lebesgue measure on $([0, \infty), \mathcal{B})$. Then $U_f \in \mathcal{S} \otimes \mathcal{B}$ and

$$(\mu \times \lambda)(U_f) = \int_X f \, d\mu.$$

Equivalently,

$$\int_X f \, d\mu = \int_0^\infty \mu(\{x \in X : \ f(x) > t\}) \, dt.$$

Exercises 5B (selected)

1. (Setup: λ Lebesgue measure on $[0, 1]$.) Compute the indicated iterated integrals for a nonnegative f on $[0, 1]^2$ and verify they agree (Tonelli/Fubini when applicable). *(Some of the integrand details were unclear in the photo; keep the computation pattern: integrate inner variable first, then outer.)*

2. Give an example of a doubly indexed collection $\{x_{m,n}\}_{m,n \in \mathbb{Z}^+}$ of nonnegative numbers such that

$$\sum_{n=1}^{\infty} x_{m,n} = 0 \text{ for every fixed } m, \qquad \text{but} \qquad \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} x_{m,n} = +\infty.$$

One standard example: define

$$x_{m,n} := \begin{cases} 1, & n = m, \\ 0, & n \neq m. \end{cases}$$

Then for each fixed m, $\sum_n x_{m,n} = 1$ (so not 0), but $\sum_m \sum_n x_{m,n} = +\infty$. *If you really want each fixed-row sum to be* 0*, then necessarily all* $x_{m,n} = 0$ *(since* $x_{m,n} \geq 0$*), forcing the double sum to be* 0*. So the statement "row sums* $= 0$*" cannot coexist with "double sum* $= \infty$*" for nonnegative arrays.*

Section 5C (continued): Lebesgue Integration on $\mathbb{R}^n$

Review: notation and topology on $\mathbb{R}^n$.

- $\mathbb{R}^n = \{(x_1, \ldots, x_n) : x_i \in \mathbb{R}\}$.
- The ℓ^∞-norm is $\|(x_1, \ldots, x_n)\|_\infty := \max\{|x_1|, \ldots, |x_n|\}$.
- For $x \in \mathbb{R}^n$ and $\delta > 0$, the open cube (ball for $\|\cdot\|_\infty$) is

$$B(x, \delta) := \{y \in \mathbb{R}^n : \ \|y - x\|_\infty < \delta\},$$

  which has side length 2δ.
- A set $G \subseteq \mathbb{R}^n$ is open iff for every $x \in G$ there exists $\delta > 0$ such that $B(x, \delta) \subseteq G$.
- A set is closed iff its complement is open.
- (Dimension bookkeeping) $\mathbb{R}^m \times \mathbb{R}^n \cong \mathbb{R}^{m+n}$ via concatenation of coordinates.

(5.36) Result (Product of open sets is open). If $G_1 \subseteq \mathbb{R}^m$ and $G_2 \subseteq \mathbb{R}^n$ are open, then $G_1 \times G_2 \subseteq \mathbb{R}^{m+n}$ is open.

(5.37) Definition (Borel σ-algebra $\mathcal{B}_n$). A *Borel subset* of $\mathbb{R}^n$ is an element of the smallest σ-algebra on $\mathbb{R}^n$ containing all open sets. This σ-algebra is denoted $\mathcal{B}_n$.

(5.38) Result (Open sets are countable unions of open cubes).

(a) A set $G \subseteq \mathbb{R}^n$ is open iff it is a countable union of open cubes in $\mathbb{R}^n$ (e.g. cubes with rational centers and rational radii in $\|\cdot\|_\infty$).

(b) $\mathcal{B}_n$ is the smallest σ-algebra on $\mathbb{R}^n$ containing all open cubes.

(5.39) Result (Product of Borel σ-algebras).

$$\mathcal{B}_m \otimes \mathcal{B}_n = \mathcal{B}_{m+n}.$$

(5.40) Definition (Lebesgue measure on $\mathbb{R}^n$). Lebesgue measure on $\mathbb{R}^n$, denoted λ_n, is defined inductively by

$$\lambda_n := \lambda_{n-1} \times \lambda_1,$$

where λ_1 is Lebesgue measure on $(\mathbb{R}, \mathcal{B}_1)$.

Iterated-integral viewpoint. Thinking of $(x, y) \in \mathbb{R}^{n-1} \times \mathbb{R}$, for $E \in \mathcal{B}_n$ one can write

$$\int_{\mathbb{R}^n} \mathbf{1}_E(x,y)\, d\lambda_n(x,y) = \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} \mathbf{1}_E(x,y)\, d\lambda_1(y) \right) d\lambda_{n-1}(x),$$

and similarly for nonnegative (or absolutely integrable) functions f on $\mathbb{R}^n$,

$$\int_{\mathbb{R}^n} f\, d\lambda_n = \int_{\mathbb{R}^{n-1}} \cdots \int_{\mathbb{R}} f(x_1, \ldots, x_n)\, dx_n \cdots dx_1,$$

with the order of integration interchangeable (Tonelli/Fubini hypotheses).

Remark. One may also define λ_n as $\lambda_j \times \lambda_k$ for any $j + k = n$; this leads to the same σ-algebra $\mathcal{B}_n$ and the same measure.

Section 5C (continued)

(5.41) Result (Measure of a dilation). Let $t > 0$. If $E \in \mathcal{B}_n$ (Borel in $\mathbb{R}^n$), then $tE \in \mathcal{B}_n$ and

$$\lambda_n(tE) = t^n \, \lambda_n(E), \qquad tE := \{tx : x \in E\}.$$

Proof sketch.

- *Borelness of tE.* Let
$$\mathcal{E} := \{E \in \mathcal{B}_n : \ tE \in \mathcal{B}_n\}.$$
Since $x \mapsto tx$ is a homeomorphism of $\mathbb{R}^n$, it maps open sets to open sets; hence all open sets lie in $\mathcal{E}$. Also $(t(E^c)) = (tE)^c$ and $t(\bigcup_k E_k) = \bigcup_k (tE_k)$, so $\mathcal{E}$ is a σ-algebra. Thus $\mathcal{E} = \mathcal{B}_n$.

- *Scaling of measure.* First verify in $\mathbb{R}$ that $\lambda_1(tI) = t\lambda_1(I)$ for intervals I, and extend to all Borel sets by standard λ-regularity/approximation (or via a monotone-class argument starting from intervals).

- *Induction on dimension.* Assuming $\lambda_{n-1}(tA) = t^{n-1}\lambda_{n-1}(A)$, check scaling on rectangles: for $A \in \mathcal{B}_{n-1}$ and $B \in \mathcal{B}_1$,
$$\lambda_n(t(A \times B)) = \lambda_n((tA) \times (tB)) = \lambda_{n-1}(tA)\, \lambda_1(tB) = t^{n-1}\lambda_{n-1}(A) \cdot t\lambda_1(B) = t^n \lambda_n(A \times B).$$
Then extend from rectangles to all Borel sets (e.g. using the monotone class theorem inside a fixed open cube and then approximating general Borel sets by intersections with an increasing sequence of cubes).

(5.43) Definition (Open unit ball in $\mathbb{R}^n$). The (Euclidean) open unit ball in $\mathbb{R}^n$ is denoted B_n and defined by

$$B_n := \Big\{ (x_1, \ldots, x_n) \in \mathbb{R}^n : \ x_1^2 + \cdots + x_n^2 < 1 \Big\}.$$

It is open, hence $B_n \in \mathcal{B}_n$.

(5.44) Result (Volume of the unit ball).

$$\lambda_n(B_n) = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2} + 1\right)}.$$

Equivalently:

$$\lambda_{2m}(B_{2m}) = \frac{\pi^m}{m!}, \qquad \lambda_{2m+1}(B_{2m+1}) = \frac{2^{m+1}\pi^m}{1 \cdot 3 \cdot 5 \cdots (2m+1)}.$$

Values for small n.

n	$\lambda_n(B_n)$	$\approx$
1	2	2
2	π	$3.14\ldots$
3	$\frac{4\pi}{3}$	$4.19\ldots$
4	$\frac{\pi^2}{2}$	$4.93\ldots$
5	$\frac{8\pi^2}{15}$	$5.26\ldots$

From the table, $\lambda_n(B_n)$ is (at least for small n) nondecreasing in n. Also note: the smallest axis-aligned cube containing B_n is $[-1,1]^n$, so it has λ_n-measure 2^n.

(5.46) Definition (Partial derivatives). Let $G \subseteq \mathbb{R}^2$ be open and $F : G \to \mathbb{R}$. For $(x,y) \in G$, define (when the limits exist)

$$(D_1F)(x,y) := \lim_{t\to 0} \frac{F(x+t,y) - F(x,y)}{t}, \qquad (D_2F)(x,y) := \lim_{t\to 0} \frac{F(x,y+t) - F(x,y)}{t}.$$

These are the partial derivatives of F with respect to the first and second coordinates, respectively.

(5.48) Result (Equality of mixed partial derivatives / Clairaut–Schwarz). Let $G \subseteq \mathbb{R}^2$ be open and $F : G \to \mathbb{R}$. If D_1F, D_2F, $D_1(D_2F)$, and $D_2(D_1F)$ exist on G and the mixed partials are continuous on G, then

$$D_1D_2F = D_2D_1F \quad \text{on } G.$$

Example. Let $F(x,y) = x^y$ on $G = (0,\infty) \times \mathbb{R}$. Then

$$D_1F(x,y) = y\,x^{y-1}, \qquad D_2F(x,y) = x^y \ln x,$$

and

$$D_2D_1F(x,y) = \frac{\partial}{\partial y}\big(yx^{y-1}\big) = x^{y-1} + y\,x^{y-1}\ln x,$$

$$D_1D_2F(x,y) = \frac{\partial}{\partial x}\big(x^y \ln x\big) = y\,x^{y-1}\ln x + x^{y-1}.$$

Hence $D_1D_2F = D_2D_1F$ on G.

Exercises 5C (continued)

2. *Euclidean balls characterize openness.* Show that a set $G \subset \mathbb{R}^n$ is open iff for every $b = (b_1, \ldots, b_n) \in G$ there exists $r > 0$ such that

$$B(b,r) := \Big\{a = (a_1, \ldots, a_n) \in \mathbb{R}^n : \ (a_1 - b_1)^2 + \cdots + (a_n - b_n)^2 < r^2\Big\} \subseteq G.$$

Proof. ($\Rightarrow$) Assume G is open and fix $b \in G$. If $G = \mathbb{R}^n$, any $r > 0$ works. Otherwise let

$$F := \mathbb{R}^n \setminus G,$$

so F is closed and $b \notin F$. Define the distance from b to F by

$$\delta := d(b, F) := \inf_{y \in F} \|b - y\|_2.$$

Then $\delta > 0$ (if $\delta = 0$ we could find $y_k \in F$ with $y_k \to b$, and since F is closed this would force $b \in F$, a contradiction). Take $r := \delta/2$. If $a \in B(b, r)$ and $a \notin G$, then $a \in F$ and hence

$$\|a - b\|_2 \geq d(b, F) = \delta,$$

contradicting $\|a - b\|_2 < r = \delta/2$. Therefore $B(b, r) \subseteq G$.

($\Leftarrow$) Assume every $b \in G$ contains some Euclidean ball $B(b, r) \subseteq G$. To show G is open, it suffices to show $F = \mathbb{R}^n \setminus G$ is closed. Let $(x_k) \subseteq F$ with $x_k \to x$. If $x \in G$, choose $r > 0$ with $B(x, r) \subseteq G$. Then for all sufficiently large k, $x_k \in B(x, r) \subseteq G$, contradicting $x_k \in F$. Hence $x \in F$, so F is closed and G is open. $\square$

5. *Dilation and change of variables for Lebesgue measure.* Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is $\mathcal{B}_n$-measurable and $t \in \mathbb{R} \setminus \{0\}$. Define $f_t : \mathbb{R}^n \to \mathbb{R}$ by

$$f_t(x) := f(tx).$$

(a) Prove that f_t is $\mathcal{B}_n$-measurable.

(b) Prove that if $\int_{\mathbb{R}^n} f \, d\lambda_n$ is defined, then

$$\int_{\mathbb{R}^n} f_t \, d\lambda_n = \frac{1}{|t|^n} \int_{\mathbb{R}^n} f \, d\lambda_n.$$

Proof. (a) Let $S_t : \mathbb{R}^n \to \mathbb{R}^n$ be $S_t(x) = tx$. Then S_t is continuous and bijective with continuous inverse $S_{1/t}$, hence a homeomorphism and in particular Borel-measurable. Since $f_t = f \circ S_t$, and compositions of Borel maps are Borel, f_t is $\mathcal{B}_n$-measurable.

(b) First handle indicators. For $A \in \mathcal{B}_n$,

$$\int_{\mathbb{R}^n} \mathbf{1}_A(tx) \, d\lambda_n(x) = \lambda_n(\{x : \; tx \in A\}) = \lambda_n(S_t^{-1}(A)) = \lambda_n\big((1/t)A\big) = \frac{1}{|t|^n} \lambda_n(A) = \frac{1}{|t|^n} \int_{\mathbb{R}^n} \mathbf{1}_A \, d\lambda_n,$$

using the dilation formula $\lambda_n((1/t)A) = |t|^{-n} \lambda_n(A)$.

By linearity, the same holds for nonnegative simple functions $\phi = \sum_{i=1}^m a_i \mathbf{1}_{A_i}$:

$$\int \phi(tx) \, d\lambda_n(x) = \frac{1}{|t|^n} \int \phi \, d\lambda_n.$$

For $f \geq 0$ measurable, pick simple $\phi_k \uparrow f$. Then $\phi_k \circ S_t \uparrow f \circ S_t = f_t$, so by MCT,

$$\int f_t \, d\lambda_n = \lim_{k \to \infty} \int (\phi_k \circ S_t) \, d\lambda_n = \lim_{k \to \infty} \frac{1}{|t|^n} \int \phi_k \, d\lambda_n = \frac{1}{|t|^n} \int f \, d\lambda_n.$$

For general (signed) f with $\int f$ defined, apply the preceding argument to f^+ and f^- and subtract: $f = f^+ - f^-$. $\square$

Chapter 6: Banach Spaces

(6.2) Definition: Metric space

Let V be a nonempty set. A *metric* on V is a function

$$d : V \times V \to [0, \infty)$$

such that for all $f, g, h \in V$:

(i) $d(f, f) = 0$.

(ii) If $d(f, g) = 0$, then $f = g$.

(iii) $d(f, g) = d(g, f)$.

(iv) (*Triangle inequality*) $d(f, h) \leq d(f, g) + d(g, h)$.

A *metric space* is a pair (V, d) where V is a nonempty set and d is a metric on V.

Examples.

- (*Discrete metric*) On any set V,

  $$d(f, g) := \begin{cases} 0, & f = g, \\ 1, & f \neq g, \end{cases}$$

  is a metric.

- On $\mathbb{R}^n$ (with $n \in \mathbb{Z}^+$), the sup metric (a.k.a. ℓ^∞ metric):

  $$d\big((x_1, \dots, x_n), (y_1, \dots, y_n)\big) := \max\{|x_1 - y_1|, \dots, |x_n - y_n|\}$$

  is a metric.

(6.3) Definition: Open ball (and closed ball)

Let (V, d) be a metric space, $f \in V$, and $r > 0$.

- The *open ball* centered at f with radius r is

  $$B(f, r) := \{g \in V : \; d(f, g) < r\}.$$

- The *closed ball* centered at f with radius r is

  $$\overline{B}(f, r) := \{g \in V : \; d(f, g) \leq r\}.$$

(6.6)–(6.9) Definitions: Closed set, closure, limit

(i) (*Closed subset*) A subset $E \subseteq V$ is *closed* if its complement $V \setminus E$ is open.

(ii) (*Closure*) For $E \subseteq V$, the *closure* of E, denoted $\overline{E}$, is

$$\overline{E} := \Big\{ g \in V : \ B(g, \varepsilon) \cap E \neq \varnothing \text{ for every } \varepsilon > 0 \Big\}.$$

(iii) (*Limit in V*) If (f_k) is a sequence in V and $f \in V$, we say $f_k \to f$ if

$$\lim_{k \to \infty} d(f_k, f) = 0.$$

(6.9) Result: Properties/characterizations of closure

Let (V, d) be a metric space and $E \subseteq V$. Then:

(a) $\overline{E} = \{ g \in V : \ \exists (f_k) \subseteq E \text{ with } f_k \to g \}$.

(b) $\overline{E}$ is the intersection of all closed subsets of V that contain E.

(c) $\overline{E}$ is closed.

(d) E is closed $\iff$ $\overline{E} = E$.

(e) E is closed $\iff$ E contains the limit of every convergent sequence of elements of E.

In words: the closure of E is the collection of all limits of sequences from E, and E is closed iff it equals its closure.

Section 6A (continued): Continuity, Cauchy sequences, completeness

(6.10) Definition: Continuity

Let (V, d_V) and (W, d_W) be metric spaces and $T : V \to W$.

- T is *continuous at* $f \in V$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$d_V(f, g) < \delta \quad \Longrightarrow \quad d_W(T(f), T(g)) < \varepsilon.$$

- T is *continuous* if it is continuous at every $f \in V$.

(6.11) Equivalent conditions for continuity

For a function $T : V \to W$ between metric spaces, the following are equivalent:

(a) T is continuous.

(b) (*Sequential characterization*) If $f_k \to f$ in V, then $T(f_k) \to T(f)$ in W.

(c) For every open set $G \subseteq W$, the preimage $T^{-1}(G)$ is open in V.

(d) For every closed set $F \subseteq W$, the preimage $T^{-1}(F)$ is closed in V.

(6.12) Definition: Cauchy sequence

A sequence (f_k) in a metric space (V, d) is *Cauchy* if for every $\varepsilon > 0$ there exists $N \in \mathbb{Z}^+$ such that

$$d(f_j, f_k) < \varepsilon \qquad \text{for all } j, k \geq N.$$

(Interpretation: the terms eventually get arbitrarily close to each other, not necessarily to a limit that lies in V.)

(6.13) Result: Every convergent sequence is Cauchy

If $f_k \to f$ in (V, d), then (f_k) is Cauchy.

(6.14) Definition: Complete metric space

A metric space (V, d) is *complete* if every Cauchy sequence in V converges to some element of V.

Example (not complete). $(\mathbb{Q}, |\cdot|)$ is not complete: one can choose a rational Cauchy sequence (q_k) that converges in $\mathbb{R}$ to an irrational number (e.g. $\sqrt{2}$, or a non-terminating non-repeating decimal), hence it has no limit in $\mathbb{Q}$.

(6.16) Result: Connection between "complete" and "closed"

(a) A complete subset of a metric space is closed (with the subspace metric).

(b) A closed subset of a complete metric space is complete (with the subspace metric).

Remark. Every nonempty subset $E \subseteq V$ becomes a metric space with the restricted metric $d|_{E \times E}$.

Section 6B: Vector Spaces (continued)

(6.17) Recall/Definition — Complex numbers $\mathbb{C}$.

- A *complex number* is an ordered pair (a, b) with $a, b \in \mathbb{R}$, written as $a + bi$.
- The set of all complex numbers is

$$\mathbb{C} := \{a + bi : (a, b) \in \mathbb{R}^2\}.$$

- Addition and multiplication on $\mathbb{C}$:

$$(a + bi) + (c + di) = (a + c) + (b + d)i,$$

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

- If $a \in \mathbb{R}$, then $a + 0i = a$, so $\mathbb{R} \subset \mathbb{C}$. Also $0 + 1i = i$.

(6.18) Definition — Re(z), Im(z), $|z|$, limits in $\mathbb{C}$. Let $z = a + bi$ with $a, b \in \mathbb{R}$.

- The real part of z is $\mathrm{Re}(z) := a$.
- The imaginary part of z is $\mathrm{Im}(z) := b$.
- The absolute value of z is
$$|z| := \sqrt{a^2 + b^2}.$$
- If $z_1, z_2, \dots \in \mathbb{C}$ and $L \in \mathbb{C}$, then
$$\lim_{k\to\infty} z_k = L \quad \Longrightarrow \quad \lim_{k\to\infty} |z_k - L| = 0.$$

(6.19) Definition — Measurable complex-valued function. Suppose $(X, \mathcal{S})$ is a measurable space. A function $f : X \to \mathbb{C}$ is called *$\mathcal{S}$-measurable* if $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$ are both $\mathcal{S}$-measurable (real-valued) functions.

(6.20) Result — $|f|^p$ is measurable. Suppose $(X, \mathcal{S})$ is a measurable space, $f : X \to \mathbb{C}$ is $\mathcal{S}$-measurable, and $0 < p < \infty$. Then $|f|^p$ is $\mathcal{S}$-measurable.

(6.21) Definition — Integral of a complex-valued function. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f : X \to \mathbb{C}$ is $\mathcal{S}$-measurable with
$$\int |f|\, d\mu < \infty \qquad (f \in L^1(\mu)).$$
Then define
$$\int f\, d\mu \;:=\; \int \mathrm{Re}(f)\, d\mu \;+\; i \int \mathrm{Im}(f)\, d\mu.$$

(6.22) Result — Bound on the absolute value of an integral. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f : X \to \mathbb{C}$ is $\mathcal{S}$-measurable with $\int |f|\, d\mu < \infty$. Then
$$\left| \int f\, d\mu \right| \le \int |f|\, d\mu.$$

Proof (as in notes). If $\int f\, d\mu = 0$ the claim is immediate. Otherwise set
$$\alpha := \frac{\overline{\int f\, d\mu}}{\left|\int f\, d\mu\right|}, \qquad \text{so that } |\alpha| = 1.$$
Then
$$\left| \int f\, d\mu \right| = \alpha \int f\, d\mu = \int \alpha f\, d\mu = \int \mathrm{Re}(\alpha f)\, d\mu \;+\; i \int \mathrm{Im}(\alpha f)\, d\mu.$$
Taking real parts gives
$$\left| \int f\, d\mu \right| = \int \mathrm{Re}(\alpha f)\, d\mu \le \int \big|\mathrm{Re}(\alpha f)\big|\, d\mu \le \int |\alpha f|\, d\mu = \int |f|\, d\mu.$$

$\square$

(6.23) Definition — Complex conjugate. Suppose $z \in \mathbb{C}$. The *complex conjugate* of z, denoted $\overline{z}$, is defined by

$$\overline{z} := \mathrm{Re}(z) - i\ \mathrm{Im}(z).$$

Properties of complex conjugation (for $w, z \in \mathbb{C}$).

1. $z\,\overline{z} = |z|^2$.
2. $z + \overline{z} = 2\,\mathrm{Re}(z), \qquad z - \overline{z} = 2i\ \mathrm{Im}(z)$.
3. $\overline{w+z} = \overline{w} + \overline{z}, \qquad \overline{wz} = \overline{w}\,\overline{z}$.
4. $\overline{\overline{z}} = z$.
5. $|\overline{z}| = |z|$.
6. If $f \in L^1(\mu)$, then
$$\overline{\int f\,d\mu} = \int \overline{f}\,d\mu.$$

From now on, $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

Section 6C: Normed Vector Spaces

(6.33) Definition — Norm; normed vector space. A *norm* on a vector space V over $\mathbb{F}$ is a function $\|\cdot\| : V \to [0, \infty)$ such that, for all $f, g \in V$ and $\alpha \in \mathbb{F}$,

1. $\|f\| = 0 \iff f = 0$ (positive definite)
2. $\|\alpha f\| = |\alpha|\,\|f\|$ (homogeneity)
3. $\|f + g\| \leq \|f\| + \|g\|$ (triangle inequality)

A *normed vector space* is a pair $(V, \|\cdot\|)$ where V is a vector space and $\|\cdot\|$ is a norm on V.

Examples.

1. If $n \in \mathbb{Z}^+$, on $\mathbb{F}^n$ define
$$\|(a_1, \dots, a_n)\|_1 := |a_1| + \cdots + |a_n|, \qquad \|(a_1, \dots, a_n)\|_\infty := \max\{|a_1|, \dots, |a_n|\}.$$
Then $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are norms on $\mathbb{F}^n$.
2. On ℓ^2, define
$$\|(a_1, a_2, \dots)\|_2 := \left(\sum_{k=1}^{\infty} |a_k|^2\right)^{1/2}.$$
Then $\|\cdot\|_2$ is a norm on ℓ^2.

(6.36) Result — Normed vector spaces are metric spaces. Suppose $(V, \|\cdot\|)$ is a normed vector space. Define $d : V \times V \to [0, \infty)$ by

$$d(f, g) := \|f - g\|.$$

Then d is a metric on V.

(6.37) Definition — Banach space. A *complete* normed vector space is called a *Banach space.* Equivalently, a normed vector space V is Banach iff every Cauchy sequence in V converges to an element of V.

(6.40) Definition — Infinite sum in a normed vector space. Suppose $g_1, g_2, \dots$ is a sequence in a normed vector space. Define

$$\sum_{k=1}^{\infty} g_k := \lim_{n\to\infty} \sum_{k=1}^{n} g_k$$

if this limit exists. If it exists, the infinite series is said to *converge.*

(6.41) Result — Absolute convergence of series $\Longleftrightarrow$ Banach. Suppose V is a normed vector space. Then V is a Banach space iff for every sequence $(g_k)_{k\geq 1} \subset V$,

$$\sum_{k=1}^{\infty} \|g_k\| < \infty \quad \Longrightarrow \quad \sum_{k=1}^{\infty} g_k \text{ converges in } V.$$

(6.43) Definition — Bounded linear map; operator norm; $B(V, W)$. Suppose V, W are normed vector spaces and $T : V \to W$ is linear.

1. The *norm* of T is
$$\|T\| := \sup\{\|Tf\| : f \in V,\ \|f\| \leq 1\}.$$
2. T is called *bounded* if $\|T\| < \infty$.
3. The set of bounded linear maps from V to W is denoted $B(V, W)$.

(6.44) Result — $\|\cdot\|$ is a norm on $B(V, W)$. Suppose V, W are normed vector spaces. Then for all $S, T \in B(V, W)$ and all $\alpha \in \mathbb{F}$,

$$\|S + T\| \leq \|S\| + \|T\|, \qquad \|\alpha T\| = |\alpha|\, \|T\|.$$

Moreover, $\|\cdot\|$ is a norm on $B(V, W)$ (so $B(V, W)$ is a normed vector space).

Proof (sketch as in notes). For $\|S + T\|$:

$$\|S+T\| = \sup_{\|f\|\leq 1} \|(S+T)f\| \leq \sup_{\|f\|\leq 1} \big(\|Sf\| + \|Tf\|\big) \leq \sup_{\|f\|\leq 1} \|Sf\| + \sup_{\|f\|\leq 1} \|Tf\| = \|S\| + \|T\|.$$

The homogeneity is immediate from linearity. □

(6.47) Result — $B(V, W)$ is Banach if W is Banach. If V is a normed vector space and W is a Banach space, then $B(V, W)$ is a Banach space.

(6.48) Result — Linear map is continuous iff bounded. A linear map between normed vector spaces is continuous iff it is bounded. (So for linear maps: continuity $\iff$ boundedness.)

Proof (as in notes). Let $T : V \to W$ be linear.

- If T is not bounded, then there exist $f_1, f_2, \dots \in V$ with $\|f_n\| \le 1$ and $\|Tf_n\| \to \infty$. Hence $\left\| \dfrac{f_n}{\|Tf_n\|} \right\| \to 0$ but
$$T\left(\frac{f_n}{\|Tf_n\|}\right) = \frac{Tf_n}{\|Tf_n\|} \not\to 0,$$
so T is not continuous at 0, hence not continuous.

- If T is bounded and $f_n \to f$ in V, then
$$\|Tf_n - Tf\| = \|T(f_n - f)\| \le \|T\|\,\|f_n - f\| \to 0,$$
so $Tf_n \to Tf$ and T is continuous.

$\square$

Exercises 6C.

(4) Every Cauchy sequence in a normed vector space is bounded. Let $(X, \|\cdot\|)$ be normed and $(x_n)_{n\ge1}$ be Cauchy. Take $\varepsilon = 1$. Then $\exists N$ such that $m, n \ge N \Rightarrow \|x_n - x_m\| < 1$. Setting $m = N$ gives $\|x_n - x_N\| < 1$ for all $n \ge N$, and thus
$$\|x_n\| \le \|x_n - x_N\| + \|x_N\| < 1 + \|x_N\| \quad (n \ge N).$$
Let
$$M := \max\{\|x_1\|, \dots, \|x_{N-1}\|,\ \|x_N\| + 1\}.$$
Then $\|x_n\| \le M$ for all n, so (x_n) is bounded. $\square$

(7) $(\ell^1, \|\cdot\|_\infty)$ is not Banach. Consider ℓ^1 equipped with the norm $\|(a_1, a_2, \dots)\|_\infty := \sup_k |a_k|$. Let
$$x^{(n)} = (1, \tfrac12, \tfrac13, \dots, \tfrac1n, 0, 0, \dots) \in \ell^1.$$
Then $(x^{(n)})$ is Cauchy in $\|\cdot\|_\infty$ and converges (in $\|\cdot\|_\infty$) to
$$x = (1, \tfrac12, \tfrac13, \dots),$$
but $x \notin \ell^1$ since $\sum_{k=1}^\infty \frac1k$ diverges. Hence $(\ell^1, \|\cdot\|_\infty)$ is not complete, so it is not a Banach space.

(8) $(\ell^2, \|\cdot\|_2)$ is Banach. Let $(x^{(n)}) \subset \ell^2$ be Cauchy in $\|\cdot\|_2$.

- (Boundedness) $\exists M$ such that $\|x^{(n)}\|_2 \le M$ for all n.
- For each coordinate k,
$$|x_k^{(n)} - x_k^{(m)}| \le \|x^{(n)} - x^{(m)}\|_2,$$
so $(x_k^{(n)})_n$ is Cauchy in $\mathbb{F}$ and $x_k := \lim_{n\to\infty} x_k^{(n)}$ exists.

Define $x := (x_1, x_2, \dots)$. Then, by Fatou,

$$\sum_{k=1}^{\infty} |x_k|^2 = \sum_{k=1}^{\infty} \liminf_{n\to\infty} |x_k^{(n)}|^2 \le \liminf_{n\to\infty} \sum_{k=1}^{\infty} |x_k^{(n)}|^2 \le M^2 < \infty,$$

so $x \in \ell^2$. Also, for fixed n,

$$\|x^{(n)} - x\|_2^2 = \sum_{k=1}^{\infty} \lim_{m\to\infty} |x_k^{(n)} - x_k^{(m)}|^2 \le \liminf_{m\to\infty} \sum_{k=1}^{\infty} |x_k^{(n)} - x_k^{(m)}|^2 = \liminf_{m\to\infty} \|x^{(n)} - x^{(m)}\|_2^2.$$

Since $(x^{(n)})$ is Cauchy, the right-hand side $\to 0$ as $n \to \infty$, hence $\|x^{(n)} - x\|_2 \to 0$. Therefore ℓ^2 is complete, i.e. a Banach space. □

Section 6D: Linear Functionals

(6.49) Definition — Linear functional. A *linear functional* on a vector space V is a linear map $\varphi : V \to \mathbb{F}$. Viewing $\mathbb{F}$ as a normed vector space with norm $\|z\| := |z|$, the field $\mathbb{F}$ is a Banach space.

Example. Let V be the vector space of sequences $(a_1, a_2, \dots)$ in $\mathbb{F}$ such that $a_k = 0$ for all but finitely many $k \in \mathbb{Z}^+$. Define

$$\varphi(a_1, a_2, \dots) := \sum_{k=1}^{\infty} a_k,$$

(which is a finite sum on V). Then φ is a linear functional on V.

(6.51) Definition — Null space. Suppose V, W are vector spaces and $T : V \to W$ is linear. The *null space* of T is

$$\mathrm{null}(T) := \{f \in V : Tf = 0\}.$$

If T is continuous between normed vector spaces, then $\mathrm{null}(T)$ is a closed subspace of V since

$$\mathrm{null}(T) = T^{-1}(\{0\}),$$

and $\{0\}$ is closed.

(6.52) Result — Bounded linear functionals. Suppose V is a normed vector space and $\varphi : V \to \mathbb{F}$ is a linear functional that is not identically 0. Then the following are equivalent:

1. φ is a bounded linear functional.
2. φ is a continuous linear functional.
3. $\mathrm{null}(\varphi)$ is a closed subspace of V.
4. $\mathrm{null}(\varphi) \ne V$.

(6.53) Definition — Family. A *family* $\{e_k\}_{k\in\Gamma}$ in a set V is a function $e : \Gamma \to V$, with $e(k)$ denoted by e_k.

(6.54) Definition — Linearly independent; span; basis. Suppose $\{e_k\}_{k\in\Gamma}$ is a family in a vector space V.

- The family is *linearly independent* if there do not exist a finite nonempty subset $\Omega \subset \Gamma$ and scalars $\{a_j\}_{j\in\Omega} \subset \mathbb{F}$, not all zero, such that
$$\sum_{j\in\Omega} a_j e_j = 0.$$
- The *span* of $\{e_k\}_{k\in\Gamma}$ is
$$\operatorname{span}\{e_k\}_{k\in\Gamma} := \left\{ \sum_{j\in\Omega} a_j e_j \;:\; \Omega \subset \Gamma \text{ finite},\ a_j \in \mathbb{F} \right\}.$$
- V is *finite-dimensional* if $\exists$ a finite set Γ and a family $\{e_k\}_{k\in\Gamma} \subset V$ such that $\operatorname{span}\{e_k\}_{k\in\Gamma} = V$. Otherwise V is *infinite-dimensional*.
- A family in V is a *basis* of V if it is linearly independent and its span equals V.

(6.56) Definition — Maximal element. Suppose $\mathcal{A}$ is a collection of subsets of a set V. A set $T \in \mathcal{A}$ is called a *maximal element* of $\mathcal{A}$ if there does not exist $T' \in \mathcal{A}$ such that $T \subsetneq T'$.

(6.57) Result — Bases as maximal elements. Suppose V is a vector space. Then a subset of V is a basis iff it is a maximal element of the collection of linearly independent subsets of V.

(6.58) Definition — Chain. A collection $\mathcal{C}$ of subsets of a set V is called a *chain* if for $\Omega, \Gamma \in \mathcal{C}$ one has
$$\Omega \subset \Gamma \quad \text{or} \quad \Gamma \subset \Omega.$$

(6.60) Result — Zorn's Lemma. Suppose V is a set and $\mathcal{A}$ is a collection of subsets of V with the property that for every chain $\mathcal{C} \subset \mathcal{A}$, the union $\bigcup_{S\in\mathcal{C}} S$ belongs to $\mathcal{A}$. Then $\mathcal{A}$ contains a maximal element.

(6.61) Result — Every vector space has a basis.

Every vector space has a basis.

(6.62) Result — Discontinuous linear functionals. Every infinite-dimensional normed vector space has a discontinuous linear functional.

Proof (construction as in notes). Let V be an infinite-dimensional normed vector space. Then V has a (Hamel) basis $\{e_k\}_{k\in\Gamma}$, and since $\dim V = \infty$, the index set Γ is infinite. Choose an injection $\mathbb{Z}^+ \subset \Gamma$ (i.e. identify a countably infinite subset of basis vectors). Define $\varphi : V \to \mathbb{F}$ by setting
$$\varphi(e_j) := j\,\|e_j\| \quad (j \in \mathbb{Z}^+), \qquad \varphi(e_k) := 0 \quad (k \in \Gamma \setminus \mathbb{Z}^+),$$
and extending linearly. More precisely, for any finite $\Omega \subset \Gamma$ and scalars $(a_j)_{j\in\Omega}$,
$$\varphi\left(\sum_{j\in\Omega} a_j e_j\right) := \sum_{j\in\Omega\cap\mathbb{Z}^+} a_j\, j\, \|e_j\|.$$

Then for $j \in \mathbb{Z}^+$,

$$\left\| \frac{e_j}{\|e_j\|} \right\| = 1, \qquad \varphi\left(\frac{e_j}{\|e_j\|} \right) = j \to \infty,$$

so φ is not bounded, hence not continuous. $\square$

(6.63) Result — Extension lemma (one-step Hahn–Banach). Suppose V is a *real* normed vector space, U is a subspace of V, and $\psi : U \to \mathbb{R}$ is a bounded linear functional. If $h \in V \setminus U$, then ψ can be extended to a bounded linear functional $\widetilde{\psi}$ on $U + \mathbb{R}h$ such that

$$\|\widetilde{\psi}\| = \|\psi\|.$$

Here

$$U + \mathbb{R}h := \{f + \alpha h : f \in U,\ \alpha \in \mathbb{R}\}.$$

(6.67) Definition — Graph. Suppose $T : V \to W$ is a function (from a set V to a set W). The *graph* of T is the subset of $V \times W$ defined by

$$\operatorname{graph}(T) := \{(f, T(f)) \in V \times W :\ f \in V\}.$$

(6.68) Result — Function properties in terms of graphs. Suppose V, W are normed vector spaces and $T : V \to W$ is a function.

1. T is linear $\iff$ $\operatorname{graph}(T)$ is a subspace of $V \times W$.
2. If $U \subset V$ and $S : U \to W$ is a function, then T is an extension of S

$$\iff \quad \operatorname{graph}(S) \subset \operatorname{graph}(T).$$

3. If $T : V \to W$ is linear and $c \in [0, \infty)$, then

$$\|T\| \le c \iff \|g\| \le c\|f\| \quad \forall (f, g) \in \operatorname{graph}(T).$$

(6.69) Result — Hahn–Banach theorem. Suppose V is a normed vector space, U is a subspace of V, and $\psi : U \to \mathbb{F}$ is a bounded linear functional. Then ψ can be extended to a bounded linear functional on V whose norm equals $\|\psi\|$.

(6.71) Result — Dual space. Suppose V is a normed vector space. The *dual space* of V, denoted V', is the normed vector space consisting of all bounded linear functionals on V:

$$V' := B(V, \mathbb{F}) \qquad \text{(in particular } V' = B(V, \mathbb{R}) \text{ if } \mathbb{F} = \mathbb{R}).$$

By (6.47), the dual space of every normed vector space is a Banach space.

(6.72) Result — Norming functional (Hahn–Banach corollary). For $f \in V$,

$$\|f\| = \sup\{|\varphi(f)| :\ \varphi \in V',\ \|\varphi\| \le 1\}.$$

Moreover, if $f \neq 0$, then $\exists\, \varphi \in V'$ such that $\|\varphi\| = 1$ and $|\varphi(f)| = \|f\|$ (and in the real case one may take $\varphi(f) = \|f\|$).

(6.73) Result — Condition to be in the closure of a subspace. Suppose U is a subspace of a normed vector space V and $h \in V$. Then

$$h \in \overline{U} \iff \varphi(h) = 0 \ \ \forall \varphi \in V' \ \text{ such that } \ \varphi|_U = 0.$$

Section 6E: Consequences of Baire's Theorem

(6.74) Definition — Interior. Suppose U is a subset of a metric space V. The *interior* of U, denoted $\text{int}(U)$, is the set of all $f \in U$ such that there exists an open ball in V centered at f with positive radius that is contained in U.

- The interior of an open subset of a metric space is open.
- $\text{int}(U)$ is the largest open subset of V contained in U.

(6.75) Definition — Dense. A subset U of a metric space V is called *dense* in V if $\overline{U} = V$.

- Example: $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ are both dense in $\mathbb{R}$ (with the standard metric $d(x, y) = |x - y|$).
- U is dense in $V \iff$ every nonempty open subset of V contains at least one element of U.
- U has empty interior $\iff V \setminus U$ is dense in V.

(6.76) Result — Baire's theorem. Let V be a complete metric space.

1. V is not the countable union of closed subsets of V with empty interior.
2. The countable intersection of dense open subsets of V is nonempty.

(6.80) Result. There does not exist a countable collection of closed subsets of $\mathbb{R}$ whose union equals $\mathbb{R} \setminus \mathbb{Q}$. (In particular, the set of irrational numbers is not a countable union of closed sets.)

Open Mapping Theorem. Suppose V, W are Banach spaces and $T : V \to W$ is a bounded linear map *onto* W. Then $T(G)$ is an open subset of W for every open subset $G \subset V$.

(6.83) Result — Bounded Inverse Theorem. Suppose V, W are Banach spaces and $T : V \to W$ is a one-to-one (injective) bounded linear map *from* V *onto* W. Then $T^{-1} : W \to V$ is a bounded linear map.

(In other words: if a bounded linear map between Banach spaces has an algebraic inverse, then the inverse is aut

(6.84) Result — Product of Banach Spaces. Suppose V, W are Banach spaces. Then $V \times W$ is a Banach space when equipped with the norm

$$\|(f, g)\| := \max\{\|f\|, \|g\|\} \qquad (f \in V, \ g \in W).$$

With this norm, a sequence $(f_1, g_1), (f_2, g_2), \dots \in V \times W$ converges to (f, g) iff

$$\|f_n - f\| \to 0 \quad \text{and} \quad \|g_n - g\| \to 0.$$

(6.85) Result — Closed Graph Theorem. Suppose V, W are Banach spaces and $T : V \to W$ is a function. Then

$$T \text{ is a bounded linear map} \quad \Longleftrightarrow \quad \text{graph}(T) \text{ is a closed subspace of } V \times W.$$

(6.86) Result — Principle of Uniform Boundedness (Banach–Steinhaus). Suppose V is a Banach space, W is a normed vector space, and $\mathcal{A}$ is a family of bounded linear maps $T : V \to W$ such that for every $f \in V$,

$$\sup\{\|Tf\| : \; T \in \mathcal{A}\} < \infty.$$

Then

$$\sup\{\|T\| : \; T \in \mathcal{A}\} < \infty.$$

(In words: pointwise bounded $\Rightarrow$ uniformly bounded on the unit ball.)

Proof (outline as in notes). For $n \in \mathbb{Z}^+$ define

$$V_n := \{\, f \in V : \; \|Tf\| \le n \text{ for all } T \in \mathcal{A} \,\}.$$

The hypothesis implies $V = \bigcup_{n=1}^{\infty} V_n$. Each V_n is closed (each T is continuous, and V_n is an intersection of closed sets). By Baire's theorem, $\exists\, n \in \mathbb{Z}^+$, $\exists\, h \in V$, $\exists\, r > 0$ such that $B(h, r) \subset V_n$.

Now let $g \in V$ with $\|g\| \le 1$. Then $h \in V_n$ and $h + rg \in B(h, r) \subset V_n$, hence for all $T \in \mathcal{A}$,

$$\|T(h + rg)\| \le n \quad \text{and} \quad \|Th\| \le n.$$

Therefore

$$\|Tg\| = \frac{1}{r}\|T(h + rg) - Th\| \le \frac{1}{r}\big(\|T(h + rg)\| + \|Th\|\big) \le \frac{2n}{r}.$$

Taking sup over $T \in \mathcal{A}$ and then over $\|g\| \le 1$ gives $\sup_{T \in \mathcal{A}} \|T\| \le 2n/r < \infty$. $\square$

Chapter 7: ℓ^p-spaces

Section 7A: $L^p(\mu)$

(7.1) Definition — p-norm $\|f\|_p$; $\|f\|_\infty$. Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $0 < p < \infty$, and $f : X \to \mathbb{F}$ is $\mathcal{S}$-measurable. Define the *p-norm* of f by

$$\|f\|_p := \left(\int |f|^p \, d\mu \right)^{1/p}.$$

Also, $\|f\|_\infty$ (the *essential supremum* of f) is defined by

$$\|f\|_\infty := \inf\big\{t > 0 : \; \mu(\{x \in X : \; |f(x)| > t\}) = 0\big\}.$$

Example (counting measure). Suppose μ is counting measure on $\mathbb{Z}^+$. If $a = (a_1, a_2, \dots)$ is a sequence in $\mathbb{F}$ and $0 < p < \infty$, then

$$\|a\|_p = \left(\sum_{k=1}^{\infty} |a_k|^p \right)^{1/p}, \qquad \|a\|_\infty = \sup\{|a_k| : \; k \in \mathbb{Z}^+\}.$$

(7.3) Definition — $L^p(\mu)$ (Lebesgue space). Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $0 < p < \infty$. The Lebesgue space $L^p(\mu)$ (or $L^p(X, \mathcal{S}, \mu)$) is defined by

$$L^p(\mu) := \{\, f : X \to \mathbb{F} \;\; \mathcal{S}\text{-measurable} : \; \|f\|_p < \infty \,\}.$$

Example: ℓ^p. When μ is counting measure on $\mathbb{Z}^+$, $L^p(\mu)$ is often denoted ℓ^p. Thus for $0 < p < \infty$,

$$\ell^p = \Big\{(a_1, a_2, \dots) : \ a_k \in \mathbb{F}, \ \sum_{k=1}^{\infty} |a_k|^p < \infty\Big\},$$

and

$$\ell^\infty = \Big\{(a_1, a_2, \dots) : \ a_k \in \mathbb{F}, \ \sup_{k \in \mathbb{Z}^+} |a_k| < \infty\Big\}.$$

(7.5) Result — $L^p(\mu)$ is a vector space. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $0 < p < \infty$. Then:

1. If $f, g \in L^p(\mu)$, then $f + g \in L^p(\mu)$ (in particular one can bound $\|f + g\|_p$ by a constant multiple of $\|f\|_p + \|g\|_p$).
2. $\|\alpha f\|_p = |\alpha| \, \|f\|_p$ for all $\alpha \in \mathbb{F}$ and $f \in L^p(\mu)$.

Hence, with the usual pointwise addition and scalar multiplication of functions, $L^p(\mu)$ is a vector space.

(7.6) Definition — Dual (conjugate) exponent p'. For $1 \le p \le \infty$, the *dual exponent* (a.k.a. conjugate exponent/index) p' is the element of $[1, \infty]$ such that

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

Examples:

$$1' = \infty, \quad \infty' = 1, \quad 2' = 2, \quad 4' = \frac{4}{3}, \quad \left(\frac{4}{3}\right)' = 4.$$

(7.8) Result — Young's inequality. Suppose $1 < p < \infty$. Then for all $a, b \ge 0$,

$$ab \le \frac{a^p}{p} + \frac{b^{p'}}{p'}.$$

(7.9) Result — Hölder's inequality. Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $1 \le p \le \infty$, and $f, h : X \to \mathbb{F}$ are $\mathcal{S}$-measurable. Then

$$\|fh\|_1 \le \|f\|_p \, \|h\|_{p'}.$$

(7.10) Result — $L^q(\mu) \subset L^p(\mu)$ if $p \le q$ and $\mu(X) < \infty$. Suppose $(X, \mathcal{S}, \mu)$ is a finite measure space and $0 < p \le q < \infty$. Then for every $f \in L^q(\mu)$,

$$\|f\|_p \le \mu(X)^{\frac{1}{p} - \frac{1}{q}} \, \|f\|_q.$$

In particular, $L^q(\mu) \subset L^p(\mu)$.

Proof. Let $f \in L^q(\mu)$ and set $r := \frac{q}{p} \ge 1$, so $r' = \frac{q}{q-p}$ (when $p < q$; the case $p = q$ is trivial). Apply Hölder to $|f|^p = (|f|^q)^{p/q} \cdot 1$ with exponents r, r':

$$\int |f|^p \, d\mu \le \left(\int \left((|f|^q)^{p/q}\right)^r d\mu\right)^{1/r} \left(\int 1^{r'} \, d\mu\right)^{1/r'} = \left(\int |f|^q \, d\mu\right)^{p/q} \mu(X)^{1-p/q}.$$

Taking pth roots gives $\|f\|_p \le \mu(X)^{\frac{1}{p} - \frac{1}{q}} \|f\|_q$. ∎

(7.12) Result — Dual formula for $\|f\|_p$. Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $1 \le p < \infty$, and $f \in L^p(\mu)$. Then

$$\|f\|_p = \sup\left\{\left|\int f\,h\,d\mu\right| :\ h \in L^{p'}(\mu),\ \|h\|_{p'} \le 1\right\}.$$

Proof. If $\|f\|_p = 0$ this is trivial. For $\|f\|_p \ne 0$, Hölder gives for any $h \in L^{p'}$ with $\|h\|_{p'} \le 1$,

$$\left|\int fh\,d\mu\right| \le \int |fh|\,d\mu \le \|f\|_p\|h\|_{p'} \le \|f\|_p,$$

so the supremum is $\le \|f\|_p$.

For the reverse inequality (when $1 < p < \infty$), define

$$h(x) := \begin{cases} \dfrac{\overline{f(x)}\,|f(x)|^{p-2}}{\|f\|_p^{p-1}}, & f(x) \ne 0, \\ 0, & f(x) = 0. \end{cases}$$

Then $\|h\|_{p'} = 1$ and

$$\int fh\,d\mu = \frac{1}{\|f\|_p^{p-1}} \int |f|^p\,d\mu = \|f\|_p.$$

Hence the supremum is $\ge \|f\|_p$. $\square$

(7.14) Result — Minkowski's inequality. Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $1 \le p \le \infty$, and $f, g \in L^p(\mu)$. Then

$$\|f + g\|_p \le \|f\|_p + \|g\|_p.$$

Proof (for $1 \le p < \infty$). Let $h \in L^{p'}(\mu)$ with $\|h\|_{p'} \le 1$. Then

$$\left|\int (f+g)h\,d\mu\right| \le \left|\int fh\,d\mu\right| + \left|\int gh\,d\mu\right| \le \|f\|_p + \|g\|_p.$$

Taking the supremum over such h and using (7.12) yields $\|f+g\|_p \le \|f\|_p + \|g\|_p$. (The case $p = \infty$ is handled separately via essential sup.) $\square$

Exercises 7A

(2) $\|\cdot\|_\infty$ properties and $L^\infty(\mu)$ is a vector space. Prove that

$$\|f + g\|_\infty \le \|f\|_\infty + \|g\|_\infty, \qquad \|\alpha f\|_\infty = |\alpha|\,\|f\|_\infty \quad (f, g \in L^\infty(\mu),\ \alpha \in \mathbb{F}),$$

and conclude that $L^\infty(\mu)$ is a vector space.

Proof (as in notes). Let $\varepsilon > 0$. By definition of essential supremum, there exist null sets N_f, N_g such that on $(N_f \cup N_g)^c$,

$$|f| \le \|f\|_\infty + \varepsilon, \qquad |g| \le \|g\|_\infty + \varepsilon.$$

Hence on $(N_f \cup N_g)^c$,

$$|f + g| \le |f| + |g| \le \|f\|_\infty + \|g\|_\infty + 2\varepsilon,$$

so $\|f + g\|_\infty \le \|f\|_\infty + \|g\|_\infty + 2\varepsilon$. Letting $\varepsilon \downarrow 0$ gives $\|f + g\|_\infty \le \|f\|_\infty + \|g\|_\infty$.

For homogeneity: if $\alpha = 0$ it is trivial. If $\alpha \ne 0$, then for any $\varepsilon > 0$, $|f| \le \|f\|_\infty + \varepsilon$ a.e. implies $|\alpha f| \le |\alpha|(\|f\|_\infty + \varepsilon)$ a.e., so $\|\alpha f\|_\infty \le |\alpha|\,\|f\|_\infty$. Apply this with $f = (1/\alpha)(\alpha f)$ to get $\|f\|_\infty \le (1/|\alpha|)\|\alpha f\|_\infty$, hence equality.

Closure under addition/scalar multiplication follows, so $L^\infty(\mu)$ is a vector space. $\square$

(6) Characterizing equality in $\|fh\|_p \le \|f\|_p\|h\|_\infty$. Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $1 \le p < \infty$, $f \in L^p(\mu)$, and $h \in L^\infty(\mu)$. Prove that

$$\|fh\|_p = \|f\|_p\,\|h\|_\infty \iff |h(x)| = \|h\|_\infty \text{ for a.e. } x \in \{f \neq 0\}.$$

Proof (as in notes). We always have

$$\|fh\|_p^p = \int |f|^p|h|^p\,d\mu \le \|h\|_\infty^p \int |f|^p\,d\mu = \|h\|_\infty^p\|f\|_p^p.$$

Assume equality holds. Then

$$0 = \|h\|_\infty^p\|f\|_p^p - \|fh\|_p^p = \int |f|^p\big(\|h\|_\infty^p - |h|^p\big)\,d\mu.$$

The integrand is nonnegative, hence it must vanish a.e. Thus for a.e. x with $|f(x)| > 0$ we have $\|h\|_\infty^p - |h(x)|^p = 0$, i.e. $|h(x)| = \|h\|_\infty$.

Conversely, if $|h| = \|h\|_\infty$ a.e. on $\{f \neq 0\}$, then $|f|^p|h|^p = \|h\|_\infty^p|f|^p$ a.e., so $\|fh\|_p^p = \|h\|_\infty^p\|f\|_p^p$ and hence $\|fh\|_p = \|f\|_p\|h\|_\infty$. □

(22) Cross-sections of an L^p function on a product space. Suppose $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{T}, \nu)$ are σ-finite measure spaces and $0 < p < \infty$. If $F \in L^p(\mu \otimes \nu)$, prove that:

$$F_x(\cdot) := F(x, \cdot) \in L^p(\nu) \text{ for a.e. } x \in X, \qquad F^y(\cdot) := F(\cdot, y) \in L^p(\mu) \text{ for a.e. } y \in Y.$$

Proof (Tonelli/Fubini argument). Since $F \in L^p(\mu \otimes \nu)$, we have $|F|^p \in L^1(\mu \otimes \nu)$. Define

$$g(x) := \int_Y |F(x, y)|^p\,d\nu(y).$$

By Tonelli, g is measurable and

$$\int_X g(x)\,d\mu(x) = \int_{X\times Y} |F|^p\,d(\mu \otimes \nu) < \infty.$$

Hence $g(x) < \infty$ for a.e. x, which is exactly $F_x \in L^p(\nu)$ a.e.

Similarly define

$$h(y) := \int_X |F(x, y)|^p\,d\mu(x),$$

and Tonelli gives $\int_Y h\,d\nu = \int_{X\times Y} |F|^p\,d(\mu\otimes\nu) < \infty$, so $h(y) < \infty$ for a.e. y, i.e. $F^y \in L^p(\mu)$ a.e. □

(24) $L^p(\mathbb{R})$ functions have p-Lebesgue points. Suppose $1 \le p < \infty$ and $F \in L^p(\mathbb{R})$. Prove that

$$\lim_{t\to 0} \frac{1}{2t}\int_{b-t}^{b+t} |F(x) - F(b)|^p\,dx = 0 \quad \text{for a.e. } b \in \mathbb{R}.$$

Proof (Lebesgue differentiation as used in notes). Since $F \in L^p(\mathbb{R})$, we have $|F|^p \in L^1(\mathbb{R})$, so the Lebesgue Differentiation Theorem applies. For a.e. b, apply the differentiation theorem to the locally integrable function

$$g_b(x) := |F(x) - F(b)|^p,$$

to get

$$\lim_{t\to 0} \frac{1}{2t}\int_{b-t}^{b+t} g_b(x)\,dx = g_b(b) = |F(b) - F(b)|^p = 0.$$

Thus the desired limit holds for almost every $b \in \mathbb{R}$. □

Section 7B: $L^p(\mu)$

(7.15) Definition — $\mathcal{Z}(\mu)$; $\widetilde{f}$. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $0 < p \le \infty$.

- $\mathcal{Z}(\mu)$ denotes the set of $\mathcal{S}$-measurable functions $f : X \to \mathbb{F}$ such that $f = 0$ a.e.
- For $f \in L^p(\mu)$, let $\widetilde{f}$ be the subset of $L^p(\mu)$ defined by

$$\widetilde{f} := \{\, f + z : \ z \in \mathcal{Z}(\mu) \,\}.$$

(7.16) Definition — $L^p(\mu)$ (as equivalence classes). Suppose μ is a measure and $0 < p \le \infty$. Let $L^p(\mu)$ denote the collection of equivalence classes of L^p-functions defined by

$$L^p(\mu) := \{\widetilde{f} : \ f \in \mathcal{L}^p(\mu)\},$$

where $\mathcal{L}^p(\mu)$ denotes the set of $\mathcal{S}$-measurable functions $f : X \to \mathbb{F}$ with $\|f\|_p < \infty$.

For $\widetilde{f}, \widetilde{g} \in L^p(\mu)$ and $\alpha \in \mathbb{F}$, define

$$\widetilde{f} + \widetilde{g} := \widetilde{(f+g)}, \qquad \alpha\widetilde{f} := \widetilde{(\alpha f)}.$$

Remark. One can think of elements of $L^p(\mu)$ as equivalence classes of functions in $\mathcal{L}^p(\mu)$, where two functions are equivalent iff they agree a.e. An element of $\mathcal{L}^p(\mu)$ is a function; an element of $L^p(\mu)$ is a set of functions, any two of which agree a.e.

(7.17) Definition — $\|\cdot\|_p$ on $L^p(\mu)$. Suppose μ is a measure and $0 < p \le \infty$. Define $\|\cdot\|_p$ on $L^p(\mu)$ by

$$\|\widetilde{f}\|_p := \|f\|_p \qquad (f \in \mathcal{L}^p(\mu)).$$

This is well-defined: if $\widetilde{f} = \widetilde{g}$, then $f = g$ a.e., hence $\|f\|_p = \|g\|_p$.

(7.18) Result — $L^p(\mu)$ is a normed vector space. Suppose μ is a measure and $1 \le p \le \infty$. Then $L^p(\mu)$ is a vector space and $\|\cdot\|_p$ is a norm on $L^p(\mu)$. Moreover, $L^p(\mu)$ is a quotient space:

$$L^p(\mu) \cong \mathcal{L}^p(\mu)/\mathcal{Z}(\mu).$$

(7.19) Definition — $L^p(E)$ for $E \subset \mathbb{R}$. If E is a Borel (equivalently: Lebesgue measurable) subset of $\mathbb{R}$ and $0 < p \le \infty$, then

$$L^p(E) \text{ means } L^p(E, \lambda_E),$$

where λ_E denotes Lebesgue measure restricted to the Borel subsets of $\mathbb{R}$ that are contained in E.

(7.20) Result — Cauchy sequences in $L^p(\mu)$ converge. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $1 \le p \le \infty$. If (f_n) is a Cauchy sequence in $L^p(\mu)$, then there exists $f \in L^p(\mu)$ such that

$$\lim_{n\to\infty} \|f_n - f\|_p = 0.$$

Proof (as in notes, for $1 \le p < \infty$). It suffices to find a subsequence (f_{n_k}) and $f \in \mathcal{L}^p(\mu)$ such that $\|f_{n_k} - f\|_p \to 0$; then the full sequence converges as well.

Since (f_n) is Cauchy, choose a subsequence (f_{n_k}) such that

$$\sum_{k=1}^{\infty} \|f_{n_k} - f_{n_{k-1}}\|_p < \infty, \qquad (f_{n_0} := 0 \text{ or any fixed start}).$$

Define for each m,

$$g_m(x) := \sum_{k=1}^{m} |f_{n_k}(x) - f_{n_{k-1}}(x)|, \qquad g(x) := \sum_{k=1}^{\infty} |f_{n_k}(x) - f_{n_{k-1}}(x)|.$$

By Minkowski,

$$\|g_m\|_p \le \sum_{k=1}^{m} \|f_{n_k} - f_{n_{k-1}}\|_p,$$

so $\sup_m \|g_m\|_p < \infty$. Also $g_m \uparrow g$ pointwise, so by monotone convergence,

$$\int g^p \, d\mu = \lim_{m\to\infty} \int g_m^p \, d\mu \le \left(\sum_{k=1}^{\infty} \|f_{n_k} - f_{n_{k-1}}\|_p \right)^p < \infty,$$

hence $g(x) < \infty$ for a.e. x.

For those x with $g(x) < \infty$, the real series $\sum_{k\ge1}(f_{n_k}(x) - f_{n_{k-1}}(x))$ converges absolutely, so define

$$f(x) := \sum_{k=1}^{\infty} (f_{n_k}(x) - f_{n_{k-1}}(x)) = \lim_{m\to\infty} f_{n_m}(x) \quad \text{for a.e. } x,$$

and set $f(x) = 0$ on the null set where the limit is not defined.

Then $f \in \mathcal{L}^p(\mu)$ and one can show $\|f_{n_k} - f\|_p \to 0$ (using Fatou and the fact that $|f_{n_k} - f| \le \sum_{j>k} |f_{n_j} - f_{n_{j-1}}|$ a.e.). Thus the subsequence converges in L^p, and hence (f_n) converges in L^p. □

(7.23) Result — L^p convergence implies a.e. convergence along a subsequence. Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $1 \le p \le \infty$, $f \in L^p(\mu)$, and $(f_n) \subset L^p(\mu)$ satisfies

$$\lim_{n\to\infty} \|f_n - f\|_p = 0.$$

Then there exists a subsequence (f_{n_k}) such that

$$\lim_{k\to\infty} f_{n_k}(x) = f(x) \qquad \text{for a.e. } x \in X.$$

(7.24) Result — $L^p(\mu)$ is Banach. Suppose μ is a measure and $1 \le p \le \infty$. Then $L^p(\mu)$ is a Banach space.

(7.25) Result — Natural map $L^{p'}(\mu) \to (L^p(\mu))'$ preserves norms. Suppose μ is a measure and $1 \le p < \infty$. For $h \in L^{p'}(\mu)$ define $\varphi_h : (L^p(\mu)) \to \mathbb{F}$ by

$$\varphi_h(\widetilde{f}) := \int f h \, d\mu.$$

Then $h \mapsto \varphi_h$ is an injective linear map from $L^{p'}(\mu)$ into $(L^p(\mu))'$, and

$$\|\varphi_h\| = \|h\|_{p'} \qquad \forall\, h \in L^{p'}(\mu).$$

(7.26) Result — Dual space of ℓ^p identified with $\ell^{p'}$. Suppose $1 \le p < \infty$. For $b = (b_1, b_2, \dots) \in \ell^{p'}$ define $\varphi_b : \ell^p \to \mathbb{F}$ by

$$\varphi_b(a) := \sum_{k=1}^{\infty} a_k b_k, \qquad a = (a_1, a_2, \dots) \in \ell^p.$$

Then $b \mapsto \varphi_b$ is an injective linear map from $\ell^{p'}$ onto $(\ell^p)'$, and

$$\|\varphi_b\| = \|b\|_{p'} \qquad \forall\, b \in \ell^{p'}.$$

Exercises 7B

(2) If $0 < p < 1$, then $\|\cdot\|$ is not a norm on $\mathbb{F}^n$. Suppose $n \ge 1$ and $0 < p < 1$. Define $\|\cdot\|$ on $\mathbb{F}^n$ by

$$\|(a_1, \dots, a_n)\| := \big(|a_1|^p + \cdots + |a_n|^p\big)^{1/p}.$$

Then $\|\cdot\|$ is *not* a norm.

Proof. Positive definiteness and homogeneity hold:

$$\|a\| = 0 \iff a = 0, \qquad \|\alpha a\| = |\alpha|\|a\|.$$

The triangle inequality fails. Let $u = (1, 0, \dots, 0)$ and $v = (0, 1, 0, \dots, 0)$. Then

$$\|u\| = \|v\| = 1, \qquad \|u + v\| = (1^p + 1^p)^{1/p} = 2^{1/p}.$$

Since $0 < p < 1$, we have $1/p > 1$, so $2^{1/p} > 2$. Hence

$$\|u + v\| > \|u\| + \|v\|,$$

contradicting the triangle inequality. □

(3a) $L^p(\mathbb{R})$ is separable for $1 \le p < \infty$. Prove there is a countable subset of $L^p(\mathbb{R})$ whose closure equals $L^p(\mathbb{R})$.

Sketch (as in notes). Let $\mathcal{I}$ be the family of all finite unions of half-open intervals $(a, b]$ with $a, b \in \mathbb{Q}$, $a < b$; then $\mathcal{I}$ is countable. Let $\mathcal{D}$ be the set of simple functions of the form

$$s = \sum_{j=1}^{m} q_j \mathbf{1}_{A_j},$$

where $m \in \mathbb{Z}^+$, $q_j \in \mathbb{Q}$, and $A_j \in \mathcal{I}$ are pairwise disjoint. Then $\mathcal{D}$ is countable. Using (i) density of simple functions in L^p for $p < \infty$, (ii) approximation of measurable sets by finite unions of rational half-open intervals (regularity of Lebesgue measure on $\mathbb{R}$), and (iii) approximation of coefficients by rationals, one shows: for every $f \in L^p(\mathbb{R})$ and every $\varepsilon > 0$, there exists $s \in \mathcal{D}$ with $\|f - s\|_p < \varepsilon$. Thus $\overline{\mathcal{D}} = L^p(\mathbb{R})$. □

(3b) $L^\infty(\mathbb{R})$ is not separable. Prove there is *no* countable subset of $L^\infty(\mathbb{R})$ whose closure equals $L^\infty(\mathbb{R})$.

Proof (as in notes). For $k \in \mathbb{Z}^+$ let

$$I_k := (2^{-(k+1)},\, 2^{-k}) \subset (0,1),$$

so the sets I_k are pairwise disjoint and have positive measure. For each subset $S \subset \mathbb{Z}^+$ define

$$f_S := \sum_{k \in S} \mathbf{1}_{I_k} \in L^\infty(\mathbb{R}).$$

If $S \neq T$, pick $k \in S \triangle T$. Then on I_k we have $|f_S - f_T| = 1$, hence

$$\|f_S - f_T\|_\infty = 1.$$

So $\{f_S : S \subset \mathbb{Z}^+\}$ is an uncountable 1-separated subset of $L^\infty(\mathbb{R})$. But in a separable metric space, every ε-separated set is at most countable. Contradiction. $\square$

(8) Cauchy sequences in $L^\infty(\mu)$ converge. Let $(X, \mathcal{S}, \mu)$ be a measure space. If $(f_n) \subset L^\infty(\mu)$ is Cauchy in $\|\cdot\|_\infty$, then there exists $f \in L^\infty(\mu)$ such that $\|f_n - f\|_\infty \to 0$.

Proof (subsequence argument as in notes). Since (f_n) is Cauchy, choose a subsequence (f_{n_m}) such that

$$\|f_{n_{m+1}} - f_{n_m}\|_\infty \le 2^{-m} \qquad (m \ge 1).$$

Define

$$E_m := \big\{x \in X :\ |f_{n_{m+1}}(x) - f_{n_m}(x)| > 2^{-m}\big\}.$$

Because $\|\cdot\|_\infty$ is an essential supremum, $\mu(E_m) = 0$ for each m. Let $E := \bigcup_{m \ge 1} E_m$, so $\mu(E) = 0$.

On $X \setminus E$, we have $|f_{n_{m+1}}(x) - f_{n_m}(x)| \le 2^{-m}$ for all m, hence $(f_{n_m}(x))$ is Cauchy in $\mathbb{F}$ and converges pointwise. Define

$$f(x) := \lim_{m \to \infty} f_{n_m}(x) \ \ (x \in X \setminus E), \qquad f(x) := 0 \ \ (x \in E).$$

Then f is measurable, and on $X \setminus E$,

$$|f(x) - f_{n_m}(x)| \le \sum_{j=m}^{\infty} |f_{n_{j+1}}(x) - f_{n_j}(x)| \le \sum_{j=m}^{\infty} 2^{-j} = 2^{-m+1}.$$

Thus $\|f - f_{n_m}\|_\infty \le 2^{-m+1} \to 0$. Finally, since (f_n) is Cauchy, for any $\varepsilon > 0$ choose m with $2^{-m+1} < \varepsilon/2$, and then choose N such that $\|f_n - f_{n_m}\|_\infty < \varepsilon/2$ for $n \ge N$. Then for $n \ge N$,

$$\|f_n - f\|_\infty \le \|f_n - f_{n_m}\|_\infty + \|f_{n_m} - f\|_\infty < \varepsilon.$$

Hence $f_n \to f$ in $L^\infty(\mu)$. $\square$

(15) The space c_0 and its dual. Let

$$c_0 := \{a = (a_1, a_2, \dots) \in \ell^\infty :\ \lim_{k \to \infty} a_k = 0\},$$

with the norm inherited from ℓ^∞: $\|a\|_\infty = \sup_k |a_k|$.

(15a) c_0 is a Banach space. *Proof (closed subspace argument).* Let $a^{(n)} \in c_0$ and suppose $a^{(n)} \to a$ in ℓ^∞, i.e. $\|a^{(n)} - a\|_\infty \to 0$. Fix $\varepsilon > 0$. Choose n so that $\|a - a^{(n)}\|_\infty < \varepsilon/2$. Since $a^{(n)} \in c_0$, choose K so that $|a_k^{(n)}| < \varepsilon/2$ for all $k \geq K$. Then for $k \geq K$,

$$|a_k| \leq |a_k - a_k^{(n)}| + |a_k^{(n)}| < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

so $a_k \to 0$ and $a \in c_0$. Thus c_0 is closed in ℓ^∞. Since ℓ^∞ is Banach, c_0 is Banach. $\square$

(15b) $(c_0)' \cong \ell^1$ isometrically. For $y = (y_k) \in \ell^1$, define $\Phi_y : c_0 \to \mathbb{F}$ by

$$\Phi_y(x) := \sum_{k=1}^{\infty} x_k y_k.$$

Then Φ_y is bounded and $\|\Phi_y\| = \|y\|_1$.

Conversely, if $\Lambda \in (c_0)'$, define $y_k := \Lambda(e^{(k)})$ where $e^{(k)}$ is the kth standard basis vector. For each N, let $x^{(N)} := \sum_{k=1}^{N} \operatorname{sgn}(y_k) e^{(k)} \in c_0$ (with $\|x^{(N)}\|_\infty = 1$). Then

$$\sum_{k=1}^{N} |y_k| = \Lambda(x^{(N)}) \leq \|\Lambda\| \, \|x^{(N)}\|_\infty = \|\Lambda\|.$$

So $(\sum_{k=1}^{N} |y_k|)$ is bounded, hence $y \in \ell^1$ and $\|y\|_1 \leq \|\Lambda\|$. For $x \in c_{00}$ (finite support), $\Lambda(x) = \sum_k x_k y_k$ by linearity; density of c_{00} in c_0 then yields $\Lambda = \Phi_y$ on all of c_0. Thus $y \mapsto \Phi_y$ is a surjective linear isometry $\ell^1 \to (c_0)'$. $\square$

Chapter 8: Hilbert Spaces

Section 8A: Inner Product Spaces

(8.1) Definition — Inner product; inner product space. An *inner product* on a vector space V over $\mathbb{F}$ (where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$) is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ such that for all $f, g, h \in V$ and $\alpha \in \mathbb{F}$:

1. (Positivity) $\langle f, f \rangle \in [0, \infty)$.
2. (Definiteness) $\langle f, f \rangle = 0 \iff f = 0$.
3. (Linearity in first slot) $\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle$ and $\langle \alpha f, g \rangle = \alpha \langle f, g \rangle$.
4. (Conjugate symmetry) $\langle f, g \rangle = \overline{\langle g, f \rangle}$.

A vector space equipped with an inner product is called an *inner product space.* (If $\mathbb{F} = \mathbb{R}$, the conjugate can be ignored, so $\langle f, g \rangle = \langle g, f \rangle$.)

Examples.

1. On $\mathbb{F}^n$:
$$\langle (a_1, \ldots, a_n), (b_1, \ldots, b_n) \rangle := a_1 \overline{b_1} + \cdots + a_n \overline{b_n}.$$
2. On ℓ^2:
$$\langle (a_k)_{k \geq 1}, (b_k)_{k \geq 1} \rangle := \sum_{k=1}^{\infty} a_k \overline{b_k}.$$

(8.3) Basic properties of the inner product. If V is an inner product space, then:

1. $\langle 0, g\rangle = \langle g, 0\rangle = 0$ for all $g \in V$.
2. $\langle f, g+h\rangle = \langle f, g\rangle + \langle f, h\rangle$ for all $f, g, h \in V$.
3. $\langle f, \alpha g\rangle = \overline{\alpha}\,\langle f, g\rangle$ for all $f, g \in V$ and $\alpha \in \mathbb{F}$.

(8.4) Definition — Norm associated with an inner product. If V is an inner product space, define for $f \in V$:

$$\|f\| := \sqrt{\langle f, f\rangle}.$$

For example, on $\mathbb{F}^n$ this gives $\|(a_1, \ldots, a_n)\| = (\sum_{j=1}^n |a_j|^2)^{1/2}$, and on ℓ^2 it gives $\|(a_k)\| = (\sum_{k=1}^\infty |a_k|^2)^{1/2}$.

(8.6) Result — Homogeneity of the norm. If V is an inner product space, then $\|\alpha f\| = |\alpha|\,\|f\|$ for all $f \in V$ and $\alpha \in \mathbb{F}$.

(8.7) Definition — Orthogonal. Two elements $f, g \in V$ are *orthogonal*, written $f \perp g$, if $\langle f, g\rangle = 0$.

(8.8) Result — Pythagorean theorem. If $f \perp g$ in an inner product space, then

$$\|f+g\|^2 = \|f\|^2 + \|g\|^2.$$

Proof. Expand:

$$\|f+g\|^2 = \langle f+g, f+g\rangle = \langle f, f\rangle + \langle f, g\rangle + \langle g, f\rangle + \langle g, g\rangle = \|f\|^2 + \|g\|^2.$$

□

(8.10) Result — Orthogonal decomposition. Suppose f, g are elements of an inner product space with $g \neq 0$. Then there exists $h \in V$ such that $h \perp g$ and

$$f = \frac{\langle f, g\rangle}{\|g\|^2}\, g + h.$$

Proof. Let $h := f - \frac{\langle f,g\rangle}{\|g\|^2} g$. Then

$$\langle h, g\rangle = \langle f, g\rangle - \frac{\langle f, g\rangle}{\|g\|^2}\langle g, g\rangle = \langle f, g\rangle - \langle f, g\rangle = 0.$$

□

(8.11) Result — Cauchy–Schwarz inequality. For f, g in an inner product space,

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

with equality iff one of f, g is a scalar multiple of the other.

Proof sketch (as in notes). If $g = 0$ it is trivial. Otherwise write $f = \frac{\langle f,g\rangle}{\|g\|^2} g + h$ with $h \perp g$. Then by Pythagoras,

$$\|f\|^2 = \left\|\frac{\langle f, g\rangle}{\|g\|^2} g\right\|^2 + \|h\|^2 = \frac{|\langle f, g\rangle|^2}{\|g\|^2} + \|h\|^2 \;\geq\; \frac{|\langle f, g\rangle|^2}{\|g\|^2}.$$

Hence $|\langle f, g\rangle| \leq \|f\|\,\|g\|$. Equality holds iff $h = 0$, i.e. f is a scalar multiple of g. □

Example (C–S in L^2). If $(X, \mathcal{S}, \mu)$ is a measure space and $f, g \in L^2(\mu)$, then

$$\left|\int f\overline{g}\, d\mu\right| \leq \left(\int |f|^2\, d\mu\right)^{1/2} \left(\int |g|^2\, d\mu\right)^{1/2}.$$

(8.15) Result — Triangle inequality (for the induced norm). If f, g are elements of an inner product space, then

$$\|f + g\| \leq \|f\| + \|g\|.$$

Moreover, equality holds iff one of f, g is a nonnegative real multiple of the other.

Proof. Compute

$$\|f+g\|^2 = \langle f+g, f+g\rangle = \|f\|^2+\|g\|^2+2\Re\langle f, g\rangle \leq \|f\|^2+\|g\|^2+2|\langle f, g\rangle| \leq \|f\|^2+\|g\|^2+2\|f\|\|g\| = (\|f\|+\|g\|)^2,$$

using $\Re z \leq |z|$ and Cauchy–Schwarz. Taking square roots gives the inequality. Equality forces $\Re\langle f, g\rangle = |\langle f, g\rangle| = \|f\|\|g\|$, i.e. $\langle f, g\rangle$ is a nonnegative real number and f, g are linearly dependent, hence one is a nonnegative real multiple of the other. □

(8.14) Result — The induced function $\|\cdot\|$ is a norm. If V is an inner product space and $\|f\| = \sqrt{\langle f, f\rangle}$, then $\|\cdot\|$ is a norm on V.

(8.20) Result — Parallelogram identity. For f, g in an inner product space,

$$\|f + g\|^2 + \|f - g\|^2 = 2\|f\|^2 + 2\|g\|^2.$$

Proof. Expand both sides:

$$\|f + g\|^2 = \langle f + g, f + g\rangle = \|f\|^2 + \|g\|^2 + \langle f, g\rangle + \langle g, f\rangle,$$

$$\|f - g\|^2 = \langle f - g, f - g\rangle = \|f\|^2 + \|g\|^2 - \langle f, g\rangle - \langle g, f\rangle.$$

Adding gives $2\|f\|^2 + 2\|g\|^2$. □

Exercises 8A

(2) An inner product on $C_b(\mathbb{R}, \mathbb{F})$ using rational evaluations. Let V be the vector space of bounded continuous functions $f : \mathbb{R} \to \mathbb{F}$. Let $r_1, r_2, \dots$ be an enumeration of $\mathbb{Q}$. For $f, g \in V$ define

$$\langle f, g\rangle := \sum_{k=1}^{\infty} \frac{f(r_k)\,\overline{g(r_k)}}{2^k}.$$

Then $\langle\cdot,\cdot\rangle$ is an inner product on V.

Proof. Let $\|f\|_\infty := \sup_{x\in\mathbb{R}} |f(x)| < \infty$. Then

$$\sum_{k=1}^{\infty} \left|\frac{f(r_k)\overline{g(r_k)}}{2^k}\right| \leq \|f\|_\infty\|g\|_\infty \sum_{k=1}^{\infty} 2^{-k} = \|f\|_\infty\|g\|_\infty < \infty,$$

so the series converges absolutely.

Linearity in the first slot is termwise:

$$\langle af_1 + bf_2, g\rangle = \sum_{k=1}^{\infty} \frac{(af_1(r_k) + bf_2(r_k))\overline{g(r_k)}}{2^k} = a\langle f_1, g\rangle + b\langle f_2, g\rangle.$$

Conjugate symmetry holds since

$$\overline{\langle g, f\rangle} = \overline{\sum_{k=1}^{\infty} \frac{g(r_k)\overline{f(r_k)}}{2^k}} = \sum_{k=1}^{\infty} \frac{f(r_k)\overline{g(r_k)}}{2^k} = \langle f, g\rangle.$$

Positivity:

$$\langle f, f\rangle = \sum_{k=1}^{\infty} \frac{|f(r_k)|^2}{2^k} \geq 0.$$

Definiteness: if $\langle f, f\rangle = 0$, then every summand is 0, hence $f(r_k) = 0$ for all k, i.e. $f = 0$ on $\mathbb{Q}$. Since $\mathbb{Q}$ is dense and f is continuous, $f \equiv 0$ on $\mathbb{R}$. Conversely, if $f \not\equiv 0$, choose x_0 with $|f(x_0)| > 0$; by continuity, $\exists\, \delta > 0$ such that $|f(x)| > 0$ on $(x_0 - \delta, x_0 + \delta)$, and picking $r_k \in \mathbb{Q} \cap (x_0 - \delta, x_0 + \delta)$ yields $\langle f, f\rangle > 0$. □

(2) Identity for the Cauchy–Schwarz deficit in $L^2(\mu)$. If $f, g \in L^2(\mu)$, then

$$\|f\|_2^2\|g\|_2^2 - |\langle f, g\rangle|^2 = \frac{1}{2}\iint_{X\times X} \big|f(x)g(y) - g(x)f(y)\big|^2\, d\mu(y)\, d\mu(x),$$

where $\langle f, g\rangle := \int f\,\overline{g}\, d\mu$.

Proof. Let

$$I := \frac{1}{2}\iint |f(x)g(y) - g(x)f(y)|^2\, d\mu(y)\, d\mu(x).$$

Expand $|a - b|^2 = |a|^2 + |b|^2 - 2\Re(a\bar{b})$ with $a = f(x)g(y)$ and $b = g(x)f(y)$:

$$\begin{aligned} 2I = &\iint |f(x)|^2|g(y)|^2\, d\mu(y)\, d\mu(x) + \iint |g(x)|^2|f(y)|^2\, d\mu(y)\, d\mu(x) \\ &- 2\Re\iint f(x)g(y)\,\overline{g(x)f(y)}\, d\mu(y)\, d\mu(x). \end{aligned}$$

By Tonelli/Fubini,

$$\iint |f(x)|^2|g(y)|^2\, d\mu(y)\, d\mu(x) = \Big(\int |f|^2\, d\mu\Big)\Big(\int |g|^2\, d\mu\Big) = \|f\|_2^2\|g\|_2^2,$$

and the second term is the same. For the cross term,

$$\iint f(x)g(y)\,\overline{g(x)f(y)}\, d\mu(y)\, d\mu(x) = \Big(\int f(x)\overline{g(x)}\, d\mu(x)\Big)\Big(\int g(y)\overline{f(y)}\, d\mu(y)\Big) = \langle f, g\rangle\,\overline{\langle f, g\rangle} = |\langle f, g\rangle|^2.$$

Hence $2I = 2\|f\|_2^2\|g\|_2^2 - 2|\langle f, g\rangle|^2$, so $I = \|f\|_2^2\|g\|_2^2 - |\langle f, g\rangle|^2$. □

(17) A weighted Cauchy–Schwarz inequality on $[1, \infty)$. Let λ be Lebesgue measure on $[1, \infty)$.

(a) If $f : [1, \infty) \to [0, \infty)$ is Borel measurable, then

$$\Big(\int_1^{\infty} f(x)\, dx\Big)^2 \leq \int_1^{\infty} x^2 f(x)^2\, dx.$$

Proof. Set $u(x) := xf(x)$ and $v(x) := \frac{1}{x}$. Then $v \in L^2([1,\infty))$ and

$$\int_1^\infty v(x)^2\,dx = \int_1^\infty \frac{1}{x^2}\,dx = 1.$$

If $\int_1^\infty x^2 f(x)^2\,dx = \infty$ the inequality is trivial. Otherwise $u \in L^2$ and by Cauchy–Schwarz,

$$\left(\int_1^\infty f(x)\,dx\right)^2 = \left(\int_1^\infty u(x)v(x)\,dx\right)^2 \le \left(\int_1^\infty u(x)^2\,dx\right)\left(\int_1^\infty v(x)^2\,dx\right) = \int_1^\infty x^2 f(x)^2\,dx.$$

□

(b) Equality holds iff $f(x) = \dfrac{c}{x^2}$ a.e. on $[1,\infty)$ for some constant $c \ge 0$.

Reason. Equality in Cauchy–Schwarz holds iff u and v are linearly dependent a.e., i.e. $u(x) = c\,v(x)$ a.e. Thus $xf(x) = c \cdot \frac{1}{x}$ a.e., so $f(x) = c/x^2$ a.e. □

(19) Product of inner product spaces. Suppose $V_1, \dots, V_m$ are inner product spaces. Define an inner product on $V_1 \times \cdots \times V_m$ by

$$\big\langle (f_1,\dots,f_m),(g_1,\dots,g_m)\big\rangle := \sum_{k=1}^m \langle f_k, g_k\rangle_{V_k}.$$

Then this defines an inner product on $V_1 \times \cdots \times V_m$.

(20) Continuity of the inner product map. Let V be an inner product space. The map

$$H : V \times V \to \mathbb{F}, \qquad H(f,g) = \langle f, g\rangle$$

is continuous (in fact, jointly continuous). More quantitatively, for any $f, g, f_0, g_0 \in V$,

$$|H(f,g) - H(f_0,g_0)| = |\langle f - f_0, g\rangle + \langle f_0, g - g_0\rangle| \le \|f - f_0\|\,\|g\| + \|f_0\|\,\|g - g_0\|.$$

In particular, if $(f,g) \to (f_0,g_0)$, then $H(f,g) \to H(f_0,g_0)$.

Section 8B: Orthogonality

(8.21) Definition — Hilbert space. A *Hilbert space* is an inner product space that is a Banach space with respect to the norm induced by the inner product,

$$\|f\| = \sqrt{\langle f, f\rangle}.$$

Example. If μ is a measure, then $L^2(\mu)$ with its usual inner product $\langle f, g\rangle = \int f\overline{g}\,d\mu$ is a Hilbert space.

(Common non-example.) The space c_{00} of finitely supported sequences with inner product $\langle a, b\rangle = \sum_{k\ge1} a_k\overline{b_k}$ is an inner product space but is not complete; its completion is ℓ^2.

(8.24) Definition — Distance from a point to a set. Suppose U is a nonempty subset of a normed vector space V and $f \in V$. Define

$$\operatorname{dist}(f, U) := \inf\{\|f - g\| :\ g \in U\}.$$

Moreover,

$$\operatorname{dist}(f,U) = 0 \iff f \in \overline{U}.$$

(8.25) Definition — Convex set. A subset U of a vector space V is *convex* if it contains the line segment joining any two of its points; i.e., for all $f, g \in U$ and all $t \in [0,1]$,

$$(1-t)f + tg \in U.$$

Every subspace of a vector space is convex.

(8.28) Result — Nearest point in a nonempty closed convex set (Hilbert space). Let V be a Hilbert space, $U \subset V$ be nonempty, closed, and convex, and $f \in V$. Then there exists a *unique* $g \in U$ such that

$$\|f - g\| = \operatorname{dist}(f, U).$$

(8.34) Definition — Orthogonal projection onto a closed convex set. Suppose U is a nonempty closed convex subset of a Hilbert space V. The *orthogonal projection* (metric projection) onto U is the map $P_U : V \to V$ defined by letting $P_U(f)$ be the unique element of U closest to f:

$$P_U(f) \in U, \qquad \|f - P_U(f)\| = \operatorname{dist}(f, U).$$

(8.37) Result — Orthogonal projection onto a closed subspace. Suppose U is a closed subspace of a Hilbert space V and $f \in V$. Then:

(i) $f - P_U(f)$ is orthogonal to U, i.e. $\langle f - P_U(f), g\rangle = 0$ for all $g \in U$.

(ii) If $h \in U$ and $f - h \perp U$, then $h = P_U(f)$.

(iii) $P_U : V \to V$ is linear.

(iv) $\|P_U(f)\| \le \|f\|$, with equality iff $f \in U$.

(8.38) Definition — Orthogonal complement $U^\perp$. Suppose U is a subset of an inner product space V. The *orthogonal complement* of U, denoted $U^\perp$, is defined by

$$U^\perp := \{\, h \in V : \ \langle h, g\rangle = 0 \ \forall\, g \in U \,\}.$$

In other words, $U^\perp$ is the set of vectors orthogonal to every element of U.

Example (in ℓ^2). Let

$$U := \{(a_1, 0, a_3, 0, a_5, 0, \dots) : \ (a_k) \in \ell^2\}.$$

Then

$$U^\perp = \{(0, b_2, 0, b_4, 0, b_6, \dots) : \ (b_k) \in \ell^2\}.$$

(8.40) Properties of orthogonal complements. If $U \subset V$ is any subset of an inner product space V, then:

(a) $U^\perp$ is a closed subspace of V.

(b) $U \cap U^\perp \subset \{0\}$.

(c) If $W \subset U$, then $U^\perp \subset W^\perp$.

(d) $\overline{U} \subset (U^\perp)^\perp$.

(e) If V is a Hilbert space and U is a subspace, then

$$\overline{U} = (U^\perp)^\perp,$$

so in particular U is closed $\iff U = (U^\perp)^\perp$.

(8.42) Result — Condition for a subspace to be dense. Suppose U is a subspace of a Hilbert space V. Then

$$\overline{U} = V \iff U^{\perp} = \{0\}.$$

(8.43) Result — Orthogonal decomposition. Suppose U is a closed subspace of a Hilbert space V. Then every $f \in V$ can be written uniquely in the form

$$f = g + h, \qquad g \in U, \quad h \in U^{\perp}.$$

Moreover $g = P_U(f)$ and $h = f - P_U(f)$, and $g \perp h$.

(8.44) Definition — Identity map I. Suppose V is a vector space. The *identity map* $I : V \to V$ is defined by

$$I(f) = f \qquad \forall f \in V.$$

(8.45) Result — Range and null space of orthogonal projections. Suppose U is a closed subspace of a Hilbert space V. Then:

(i) $\text{range}(P_U) = U$ and $\text{null}(P_U) = U^{\perp}$.

(ii) $\text{range}(P_{U^{\perp}}) = U^{\perp}$ and $\text{null}(P_{U^{\perp}}) = U$.

(iii) $P_{U^{\perp}} = I - P_U$.

(8.47) Result — Riesz representation theorem. Let V be a Hilbert space over $\mathbb{F}$.

- If $h \in V$, define $\varphi_h : V \to \mathbb{F}$ by $\varphi_h(f) := \langle f, h \rangle$. Then φ_h is a bounded linear functional and

$$|\varphi_h(f)| = |\langle f, h \rangle| \le \|f\| \, \|h\| \quad \Rightarrow \quad \|\varphi_h\| \le \|h\|.$$

  In fact $\|\varphi_h\| = \|h\|$.

- Conversely, if $\varphi \in V'$ is any bounded linear functional on V, then there exists a *unique* $h \in V$ such that

$$\varphi(f) = \langle f, h \rangle \qquad \forall f \in V,$$

  and moreover $\|\varphi\| = \|h\|$.

Proof sketch (matching the notes). If $\varphi = 0$, take $h = 0$. If $\varphi \neq 0$, then $\text{null}(\varphi) \neq V$, so $(\text{null}(\varphi))^{\perp} \neq \{0\}$. Pick $g \in (\text{null}(\varphi))^{\perp}$ with $\|g\| = 1$. Let $h := \overline{\varphi(g)}\, g$ (over $\mathbb{R}$, just $h = \varphi(g)g$). Then $\|h\| = |\varphi(g)|$, and one checks that for every $f \in V$, writing $f = f_0 + \alpha g$ with $f_0 \in \text{null}(\varphi)$ gives $\varphi(f) = \alpha\varphi(g) = \langle f, h \rangle$. Uniqueness follows since $\langle f, h - \tilde{h} \rangle = 0$ for all f implies $h = \tilde{h}$. Finally, $\|\varphi\| \le \|h\|$ by Cauchy–Schwarz and $\|\varphi\| \ge |\varphi(g)| = \|h\|$, so $\|\varphi\| = \|h\|$. $\square$

Exercises 8B

(2) Disprove: the inner product space from Exercise 8A(2) is Hilbert. Let $V = C_b(\mathbb{R})$ with

$$\langle f, g\rangle = \sum_{k=1}^{\infty} \frac{f(r_k)\overline{g(r_k)}}{2^k}, \qquad \|f\|^2 = \sum_{k=1}^{\infty} \frac{|f(r_k)|^2}{2^k}.$$

Show V is not complete (hence not Hilbert).

Construction (as in notes). Define a sequence (y_k) by

$$y_k := \begin{cases} 1, & r_k \le 0, \\ 0, & r_k > 0. \end{cases}$$

For each N, choose $F_N \in C_b(\mathbb{R})$ such that $0 \le F_N \le 1$ and

$$F_N(r_k) = y_k \quad \text{for } 1 \le k \le N.$$

Then for $M \ge N$,

$$\|F_M - F_N\|^2 = \sum_{k=1}^{\infty} \frac{|F_M(r_k) - F_N(r_k)|^2}{2^k} = \sum_{k>N} \frac{|F_M(r_k) - F_N(r_k)|^2}{2^k} \le \sum_{k>N} \frac{1}{2^k} = 2^{-N}.$$

Hence (F_N) is Cauchy in $\|\cdot\|$.

If $F_N \to F$ in this norm, then for each fixed k,

$$|F_N(r_k) - F(r_k)|^2 \le 2^k \|F_N - F\|^2 \to 0,$$

so $F(r_k) = y_k$ for all k.

Pick rationals $q_n \in \mathbb{Q} \cap (-\infty, 0)$ with $q_n \to 0$ and $p_n \in \mathbb{Q} \cap (0, \infty)$ with $p_n \to 0$. Then $F(q_n) = 1$ and $F(p_n) = 0$ for all n. Continuity at 0 would force both limits to equal $F(0)$, impossible. Thus no such $F \in C_b(\mathbb{R})$ exists, so V is not complete. $\square$

(3) ℓ^2-direct sums of Hilbert spaces are Hilbert. Suppose $V_1, V_2, \ldots$ are Hilbert spaces. Define

$$V := \left\{ (f_1, f_2, \ldots) : f_k \in V_k,\ \sum_{k=1}^{\infty} \|f_k\|_{V_k}^2 < \infty \right\}.$$

For $f = (f_k)$ and $g = (g_k)$ define

$$\langle f, g\rangle := \sum_{k=1}^{\infty} \langle f_k, g_k\rangle_{V_k}.$$

Then $\langle \cdot, \cdot \rangle$ is an inner product on V and V is a Hilbert space.

Key points. By Cauchy–Schwarz in each V_k, $|\langle f_k, g_k\rangle| \le \|f_k\|\|g_k\|$, so

$$\sum_{k=1}^{\infty} |\langle f_k, g_k\rangle| \le \left(\sum_{k=1}^{\infty} \|f_k\|^2\right)^{1/2} \left(\sum_{k=1}^{\infty} \|g_k\|^2\right)^{1/2} < \infty,$$

hence the inner product is well-defined and satisfies the axioms termwise.

For completeness: if $(x^{(n)})$ is Cauchy in V, then each coordinate $(x_k^{(n)})$ is Cauchy in V_k, hence converges to some $x_k \in V_k$. Fatou gives $\sum_k \|x_k\|^2 < \infty$, so $x = (x_k) \in V$, and again using Fatou/standard estimates yields $x^{(n)} \to x$ in V. Thus V is complete. $\square$

(23a) Example: strict inequality $|\varphi(f)| < \|\varphi\|\,\|f\|$ for all $f \neq 0$. Let $V = c_0$ with $\|\cdot\| = \|\cdot\|_\infty$. Define $\varphi : V \to \mathbb{F}$ by

$$\varphi(x) := \sum_{n=1}^{\infty} 2^{-n} x_n, \qquad x = (x_n) \in c_0.$$

Then φ is bounded and $\|\varphi\| = 1$, but for every $x \neq 0$,

$$|\varphi(x)| < \|x\|_\infty\,\|\varphi\|.$$

Reason.

$$|\varphi(x)| \le \sum_{n=1}^{\infty} 2^{-n}|x_n| \le \|x\|_\infty \sum_{n=1}^{\infty} 2^{-n} = \|x\|_\infty,$$

so $\|\varphi\| \le 1$. Let $x^{(N)} = (1, \dots, 1, 0, 0, \dots) \in c_0$ (first N entries 1). Then $\|x^{(N)}\|_\infty = 1$ and

$$\varphi(x^{(N)}) = \sum_{n=1}^{N} 2^{-n} = 1 - 2^{-N} \to 1,$$

so $\|\varphi\| \ge 1$, hence $\|\varphi\| = 1$.

If equality $|\varphi(x)| = \|x\|_\infty$ held for some nonzero $x \in c_0$, we would need $|x_n| = \|x\|_\infty$ for all n with $2^{-n} > 0$, which is impossible in c_0 unless $x = 0$ (since $x_n \to 0$). Hence the inequality is strict for all $x \neq 0$. $\square$

(23b) In a Hilbert space, equality is attained. If V is a Hilbert space and $\varphi \in V'$, then by Riesz there exists $h \in V$ with $\varphi(f) = \langle f, h\rangle$ and $\|\varphi\| = \|h\|$. Taking $f = h \neq 0$ gives

$$|\varphi(h)| = |\langle h, h\rangle| = \|h\|^2 = \|\varphi\|\,\|h\|,$$

so in Hilbert spaces the operator norm is achieved (unless $\varphi = 0$).

(26) If V is infinite-dimensional Hilbert, then $B(V)$ is nonseparable. Let V be an infinite-dimensional Hilbert space and $B(V)$ the Banach space of bounded linear operators $V \to V$ (with operator norm).

Let $\{e_j\}_{j\in J}$ be an orthonormal basis of V. For each sign choice $s = (s_j)_{j\in J} \in \{\pm 1\}^J$, define $U_s \in B(V)$ by

$$U_s\Big(\sum_{j\in J} \alpha_j e_j\Big) := \sum_{j\in J} s_j\,\alpha_j e_j.$$

Then $\|U_s\| = 1$ for all s (it is an isometry). If $s \neq t$, pick j with $s_j \neq t_j$. Then

$$(U_s - U_t)e_j = (s_j - t_j)e_j = \pm 2\,e_j,$$

so $\|U_s - U_t\| \ge 2$. Also $\|U_s - U_t\| \le \|U_s\| + \|U_t\| = 2$, hence $\|U_s - U_t\| = 2$. Therefore $\{U_s : s \in \{\pm 1\}^J\}$ is a 2-separated set in $B(V)$ of cardinality $2^{|J|}$ (uncountable). A separable metric space cannot contain an uncountable ε-separated set for any $\varepsilon > 0$. Thus $B(V)$ is nonseparable. $\square$

Section 8C: Orthonormal Bases

Recall (family). A family $\{e_k\}_{k\in\Gamma}$ in a set V is a function $e : \Gamma \to V$, with value at k denoted e_k.

(8.50) Definition — Orthonormal family. A family $\{e_k\}_{k\in\Gamma}$ in an inner product space V is *orthonormal* if

$$\langle e_i, e_j\rangle = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \qquad (i, j \in \Gamma).$$

Equivalently, the vectors e_i, e_j are orthogonal for $i \neq j$, and $\|e_k\| = 1$ for all $k \in \Gamma$.

Example (Fourier system in $L^2([-\pi, \pi])$). For $k \in \mathbb{Z}$, define

$$e_k(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi}}, & k = 0, \\ \dfrac{1}{\sqrt{\pi}}\cos(kx), & k \geq 1, \\ \dfrac{1}{\sqrt{\pi}}\sin(|k|x), & k \leq -1, \end{cases} \qquad x \in [-\pi, \pi].$$

Then $\{e_k\}_{k\in\mathbb{Z}}$ is an orthonormal family in $L^2([-\pi, \pi])$.

(8.52) Result — Finite orthonormal families. Suppose Ω is a finite set and $\{e_j\}_{j\in\Omega}$ is an orthonormal family in an inner product space V. Then for any scalars $\{a_j\}_{j\in\Omega} \subset \mathbb{F}$,

$$\left\|\sum_{j\in\Omega} a_j e_j\right\|^2 = \sum_{j\in\Omega} |a_j|^2.$$

Proof.

$$\left\|\sum_{j\in\Omega} a_j e_j\right\|^2 = \left\langle \sum_{j\in\Omega} a_j e_j, \sum_{k\in\Omega} a_k e_k \right\rangle = \sum_{j,k\in\Omega} a_j \overline{a_k} \langle e_j, e_k\rangle = \sum_{j\in\Omega} |a_j|^2. \quad \square$$

(8.53) Definition — Unordered sums $\sum_{k\in\Gamma} f_k$. Suppose $\{f_k\}_{k\in\Gamma}$ is a family in a normed vector space V. We say the *unordered sum* $\sum_{k\in\Gamma} f_k$ *converges* to $f \in V$ if for every $\varepsilon > 0$ there exists a finite set $\Omega \subset \Gamma$ such that

$$\left\| f - \sum_{k\in\Omega'} f_k \right\| < \varepsilon \quad \text{for all finite } \Omega' \text{ with } \Omega \subset \Omega' \subset \Gamma.$$

If this happens, we write $f = \sum_{k\in\Gamma} f_k$.

Example (nonnegative reals). If $a_k \geq 0$ for all $k \in \Gamma$ (a family in $\mathbb{R}$), then $\sum_{k\in\Gamma} a_k$ converges iff

$$\sup \left\{ \sum_{k\in\Omega} a_k : \ \Omega \subset \Gamma \text{ finite} \right\} < \infty,$$

and in that case it equals that supremum.

(8.54) Definition/Result — Linear combinations of an orthonormal family. Suppose $\{e_k\}_{k\in\Gamma}$ is an orthonormal family in a Hilbert space V, and $\{a_k\}_{k\in\Gamma}$ is a scalar family. Then:

(i) The unordered sum $\sum_{k\in\Gamma} a_k e_k$ converges in V iff $\sum_{k\in\Gamma} |a_k|^2 < \infty$.

(ii) If $\sum_{k\in\Gamma} a_k e_k$ converges, then

$$\left\| \sum_{k\in\Gamma} a_k e_k \right\|^2 = \sum_{k\in\Gamma} |a_k|^2.$$

Proof sketch. If $\sum_{k\in\Gamma} a_k e_k = g$, then (8.52) gives $\left\| \sum_{k\in\Omega} a_k e_k \right\|^2 = \sum_{k\in\Omega} |a_k|^2 \le (\|g\| + \varepsilon)^2$ for all sufficiently large finite Ω, hence $\sup_\Omega \sum_{k\in\Omega} |a_k|^2 < \infty$ and thus $\sum_{k\in\Gamma} |a_k|^2 < \infty$.

Conversely, if $\sum_{k\in\Gamma} |a_k|^2 < \infty$, choose an increasing sequence of finite sets $\Omega_1 \subset \Omega_2 \subset \cdots$ with $\bigcup_m \Omega_m = \Gamma$ and such that $\sum_{k\in\Gamma\setminus\Omega_m} |a_k|^2 \to 0$. Let $g_m := \sum_{k\in\Omega_m} a_k e_k$. Then for $n > m$,

$$\|g_n - g_m\|^2 = \sum_{k\in\Omega_n\setminus\Omega_m} |a_k|^2 \to 0,$$

so (g_m) is Cauchy and converges in V to some g. The unordered convergence definition follows from the same tail estimate, and (ii) follows by passing to the limit in (8.52). □

(8.57) Definition/Result — Bessel's inequality. Suppose $\{e_k\}_{k\in\Gamma}$ is an orthonormal family in an inner product space V and $f \in V$. Then

$$\sum_{k\in\Gamma} |\langle f, e_k\rangle|^2 \le \|f\|^2.$$

Proof (finite version ⇒ supremum). For any finite $\Omega \subset \Gamma$,

$$\left\| f - \sum_{k\in\Omega} \langle f, e_k\rangle e_k \right\|^2 = \|f\|^2 - \sum_{k\in\Omega} |\langle f, e_k\rangle|^2 \ge 0,$$

hence $\sum_{k\in\Omega} |\langle f, e_k\rangle|^2 \le \|f\|^2$ for all finite Ω, which implies the claim. □

(8.58) Result — Closure of the span of an orthonormal family. Suppose $\{e_k\}_{k\in\Gamma}$ is an orthonormal family in a Hilbert space V. Then

$$\overline{\operatorname{span}}\{e_k : k \in \Gamma\} = \left\{ \sum_{k\in\Gamma} a_k e_k : \sum_{k\in\Gamma} |a_k|^2 < \infty \right\},$$

and for every f in this closed span,

$$f = \sum_{k\in\Gamma} \langle f, e_k\rangle e_k.$$

(8.61) Definition — Orthonormal basis. An orthonormal family $\{e_k\}_{k\in\Gamma}$ in a Hilbert space V is an *orthonormal basis* (ONB) if

$$\overline{\operatorname{span}}\{e_k : k \in \Gamma\} = V.$$

(8.63) Result — Parseval's identity. Suppose $\{e_k\}_{k\in\Gamma}$ is an orthonormal basis of a Hilbert space V and $f, g \in V$. Then:

(i) $f = \sum_{k\in\Gamma} \langle f, e_k\rangle e_k$.

(ii) $\langle f, g\rangle = \sum_{k\in\Gamma} \langle f, e_k\rangle \langle e_k, g\rangle = \sum_{k\in\Gamma} \langle f, e_k\rangle\, \overline{\langle g, e_k\rangle}$.

(iii) $\|f\|^2 = \sum_{k\in\Gamma} |\langle f, e_k\rangle|^2$.

(8.64) Definition — Separable. A normed vector space V is *separable* if it has a countable subset whose closure is V.

Example. ℓ^2 is separable because the set of finitely supported sequences with rational coordinates is countable and dense in ℓ^2.

(8.67) Result — Every separable Hilbert space has an orthonormal basis.

(8.71) Result — Orthogonal projection in terms of an ONB. Suppose U is a closed subspace of a Hilbert space V and $\{e_k\}_{k\in\Gamma}$ is an orthonormal basis of U. Then for every $f \in V$,

$$P_U(f) = \sum_{k\in\Gamma} \langle f, e_k\rangle e_k.$$

(8.74) Result — Orthonormal bases as maximal orthonormal sets. Let V be a Hilbert space and let $\mathcal{A}$ be the collection of all orthonormal subsets of V (ordered by inclusion). If $\Gamma \in \mathcal{A}$, then Γ is an orthonormal basis of V iff Γ is a maximal element of $\mathcal{A}$.

(8.75) Result — Every Hilbert space has an orthonormal basis. (Uses Zorn's lemma applied to $\mathcal{A}$.)

(8.76) Result — Riesz representation via coordinates in an ONB. Suppose φ is a bounded linear functional on a Hilbert space V, and $\{e_k\}_{k\in\Gamma}$ is an orthonormal basis of V. Let

$$c_k := \varphi(e_k) \qquad (k \in \Gamma).$$

Then $\sum_{k\in\Gamma} |c_k|^2 < \infty$, and if we define

$$h := \sum_{k\in\Gamma} \overline{c_k}\, e_k \in V,$$

we have

$$\varphi(f) = \langle f, h\rangle \qquad \forall f \in V,$$

and

$$\|\varphi\| = \|h\| = \left(\sum_{k\in\Gamma} |c_k|^2\right)^{1/2}.$$

Exercises 8C

(2) Let $\{a_k\}_{k\in\Gamma} \subset \mathbb{R}$ with $a_k \geq 0$ for all k. For $\Omega \in \mathcal{F}(\Gamma) := \{\Omega \subset \Gamma : \ |\Omega| < \infty\}$ define

$$s_\Omega := \sum_{k\in\Omega} a_k.$$

Claim: The unordered sum $\sum_{k\in\Gamma} a_k$ converges $\iff$

$$M := \sup\{s_\Omega : \ \Omega \in \mathcal{F}(\Gamma)\} < \infty,$$

and in that case $\sum_{k\in\Gamma} a_k = M$.

Proof. If the unordered sum exists and equals S, then $s_\Omega \le S$ for all finite Ω, hence $M \le S < \infty$. Conversely assume $M < \infty$. Since $a_k \ge 0$, the net $\{s_\Omega\}_{\Omega \in \mathcal{F}(\Gamma)}$ is increasing: if $\Omega \subset \Omega'$ then $s_\Omega \le s_{\Omega'} \le M$. Fix $\varepsilon > 0$. By definition of supremum, choose Ω_ε with $s_{\Omega_\varepsilon} > M - \varepsilon$. Then for any $\Omega \supset \Omega_\varepsilon$,

$$M - \varepsilon < s_{\Omega_\varepsilon} \le s_\Omega \le M \quad \Rightarrow \quad |s_\Omega - M| < \varepsilon,$$

so $s_\Omega \to M$. □

(4) Let X be a normed vector space and let $\{f_k\}_{k\in\Gamma}, \{g_k\}_{k\in\Gamma} \subset X$ be families such that the unordered sums $\sum_{k\in\Gamma} f_k$ and $\sum_{k\in\Gamma} g_k$ converge. **Claim:** $\sum_{k\in\Gamma}(f_k + g_k)$ converges unordered and

$$\sum_{k\in\Gamma}(f_k + g_k) = \sum_{k\in\Gamma} f_k + \sum_{k\in\Gamma} g_k.$$

Proof. For $\Omega \in \mathcal{F}(\Gamma)$ set

$$u_\Omega := \sum_{k\in\Omega} f_k, \qquad v_\Omega := \sum_{k\in\Omega} g_k, \qquad t_\Omega := \sum_{k\in\Omega}(f_k + g_k).$$

Then $t_\Omega = u_\Omega + v_\Omega$ for all Ω. If $u_\Omega \to F$ and $v_\Omega \to G$ in X, continuity of addition gives $t_\Omega \to F + G$, i.e. the unordered sum exists and equals $F + G$. □

(7) Let V be a normed vector space, $\{f_k\}_{k\in\mathbb{Z}^+} \subset V$, and $F \in V$. For $\Omega \in \mathcal{F} := \{\Omega \subset \mathbb{Z}^+ : |\Omega| < \infty\}$ set $s_\Omega := \sum_{k\in\Omega} f_k$. **Claim:** The unordered sum $\sum_{k\in\mathbb{Z}^+} f_k = F$ iff for every bijection (permutation) $p : \mathbb{Z}^+ \to \mathbb{Z}^+$, the series $\sum_{n=1}^\infty f_{p(n)}$ converges to F.
Proof. ($\Rightarrow$) If $s_\Omega \to F$, then for any bijection p the partial sums

$$S_n := \sum_{j=1}^n f_{p(j)} = s_{\Omega_n}, \qquad \Omega_n := \{p(1), \dots, p(n)\},$$

satisfy $S_n \to F$.

($\Leftarrow$) Assume every rearrangement converges to F. We first prove the *tail lemma*:

Lemma (tail smallness). For every $\varepsilon > 0$ there exists a finite $A \subset \mathbb{Z}^+$ such that for every finite $B \subset A^c$,

$$\left\| \sum_{k\in B} f_k \right\| < \varepsilon.$$

Proof of lemma. If false, then $\exists\, \varepsilon_0 > 0$ such that for every finite A there exists a finite $B \subset A^c$ with $\left\|\sum_{k\in B} f_k\right\| \ge \varepsilon_0$. Construct disjoint finite blocks $B_1, B_2, \dots$ inductively: pick any finite B_1 with norm $\ge \varepsilon_0$; given $B_1, \dots, B_m$, take $A = \bigcup_{j\le m} B_j$ and pick $B_{m+1} \subset A^c$ with norm $\ge \varepsilon_0$. Now define a permutation p by listing elements block-by-block. Then the rearranged partial sums jump by at least ε_0 when a block is completed, so they cannot be Cauchy, contradicting convergence of the rearrangement. □

Now fix $\varepsilon > 0$ and choose A as in the lemma with tolerance $\varepsilon/3$. Pick any bijection p and choose n_0 such that

$$\left\| \sum_{j=1}^{n_0} f_{p(j)} - F \right\| < \varepsilon/3.$$

Let $A_0 := \{p(1), \dots, p(n_0)\}$ (finite). Enlarge A if needed so that $A_0 \subset A$. For any finite $\Omega \supset A$, write $\Omega = A \cup B$ with $B \subset A^c$ finite. Then

$$\|s_\Omega - F\| \le \|s_A - F\| + \left\| \sum_{k\in B} f_k \right\| < \varepsilon/3 + \varepsilon/3 < \varepsilon,$$

so $s_\Omega \to F$. $\square$

Continued Exercises 8C

(11) Let μ be a σ-finite measure on $(X, \mathcal{S})$ and ν a σ-finite measure on $(Y, \mathcal{T})$. Assume $\{e_j\}_{j\in\Omega}$ is an orthonormal basis of $L^2(X,\mu)$ and $\{f_k\}_{k\in\Gamma}$ is an orthonormal basis of $L^2(Y,\nu)$, with Ω, Γ countable. For $j \in \Omega$, $k \in \Gamma$, define

$$g_{j,k}(x,y) := e_j(x) f_k(y).$$

Claim: $\{g_{j,k}\}_{(j,k)\in\Omega\times\Gamma}$ is an orthonormal basis of $L^2(X \times Y, \mu \times \nu)$.
Proof. For (j,k) and (j',k'),

$$\langle g_{j,k}, g_{j',k'}\rangle = \int_X \int_Y e_j(x)\overline{e_{j'}(x)}\, f_k(y)\overline{f_{k'}(y)}\, d\nu(y)\, d\mu(x) = \langle e_j, e_{j'}\rangle \langle f_k, f_{k'}\rangle = \delta_{jj'}\delta_{kk'}.$$

For completeness, suppose $h \in L^2(\mu \times \nu)$ satisfies $\langle h, g_{j,k}\rangle = 0$ for all j,k. Fix j and define

$$a_j(y) := \int_X h(x,y)\, \overline{e_j(x)}\, d\mu(x).$$

Then for every k,

$$0 = \langle h, g_{j,k}\rangle = \int_Y a_j(y)\, \overline{f_k(y)}\, d\nu(y) = \langle a_j, f_k\rangle_{L^2(\nu)}.$$

Since $\{f_k\}$ is an ONB, $a_j = 0$ a.e. for each j. Hence for a.e. y, the function $x \mapsto h(x,y)$ is orthogonal to all e_j, so $h(\cdot, y) = 0$ in $L^2(\mu)$. Thus $h = 0$ in $L^2(\mu \times \nu)$, proving completeness. $\square$

(12) Let $\{e_k\}_{k\in\Gamma}$ be an orthonormal family in a Hilbert space V and assume

$$\|f\|^2 = \sum_{k\in\Gamma} |\langle f, e_k\rangle|^2 \qquad \forall f \in V.$$

Claim: $\{e_k\}$ is an orthonormal basis of V.
Proof. Let $W := \overline{\operatorname{span}}\{e_k :\ k \in \Gamma\}$. For $f \in W^\perp$ we have $\langle f, e_k\rangle = 0$ for all k, hence

$$\|f\|^2 = \sum_k |\langle f, e_k\rangle|^2 = 0 \ \Rightarrow\ f = 0.$$

Thus $W^\perp = \{0\}$, so $W = V$. $\square$

(16) Find the polynomial g of degree ≤ 4 that minimizes

$$\int_{-1}^{1} |x^5 - g(x)|^2\, dx.$$

Solution. Let $H := \operatorname{span}\{1, x, x^2, x^3, x^4\} \subset L^2([-1,1])$. The minimizer is the orthogonal projection of x^5 onto H. By symmetry, x^5 is odd and is orthogonal to the even subspace, so the projection lies in $\operatorname{span}\{x, x^3\}$: write $g(x) = ax + bx^3$. Orthogonality conditions

$$\int_{-1}^{1} (x^5 - g(x))x\, dx = 0, \qquad \int_{-1}^{1} (x^5 - g(x))x^3\, dx = 0$$

give

$$\int_{-1}^{1} x^6\,dx - a\int_{-1}^{1} x^2\,dx - b\int_{-1}^{1} x^4\,dx = 0, \qquad \int_{-1}^{1} x^8\,dx - a\int_{-1}^{1} x^4\,dx - b\int_{-1}^{1} x^6\,dx = 0.$$

Using $\int_{-1}^{1} x^{2m}\,dx = \frac{2}{2m+1}$ yields the system

$$\frac{2}{7} - a\frac{2}{3} - b\frac{2}{5} = 0, \qquad \frac{2}{9} - a\frac{2}{5} - b\frac{2}{7} = 0,$$

so $a = -\frac{5}{21}$ and $b = \frac{10}{9}$. Therefore

$$g(x) = \frac{10}{9}x^3 - \frac{5}{21}x.$$

□

(20) Let $G \subset \mathbb{C}$ be open and nonempty. Define

$$L_a^2(G) := \left\{ f \in \mathcal{O}(G) : \int_G |f(z)|^2\,d\lambda_2(z) < \infty \right\}, \qquad \langle f, g\rangle := \int_G f(z)\overline{g(z)}\,d\lambda_2(z),$$

where λ_2 is planar Lebesgue measure. **Claim:** $L_a^2(G)$ is a Hilbert space.
Proof sketch. It is an inner product space: sesquilinearity is immediate, and if $\langle f, f\rangle = 0$ then $f = 0$ a.e., hence $f \equiv 0$ by analyticity.

To prove completeness, use the *evaluation estimate* from Cauchy + Cauchy–Schwarz: if $\overline{D(a,r)} \subset G$, then

$$|f(a)|^2 \le \frac{1}{\pi r^2}\int_{D(a,r)} |f(z)|^2\,d\lambda_2(z) \le \frac{1}{\pi r^2}\|f\|_2^2,$$

hence on any compact $K \Subset G$ there exists C_K with $\sup_{z\in K}|f(z)| \le C_K\|f\|_2$. Thus an L^2-Cauchy sequence $\{f_n\} \subset L_a^2(G)$ is locally uniformly bounded, hence a normal family; extract a subsequence converging locally uniformly to some $g \in \mathcal{O}(G)$ (Montel). Since $L^2(G)$ is complete, $f_n \to f$ in $L^2(G)$ for some $f \in L^2(G)$, and along a further subsequence $f_{n_m} \to f$ a.e.; but also $f_{n_m} \to g$ pointwise, so $f = g$ a.e. Hence $g \in L_a^2(G)$ and $f_n \to g$ in L^2. □

(21) Let $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$ and $L_a^2(\mathbb{D})$ as above (the *Bergman space*).
(a) ONB. For $m, n \ge 0$,

$$\langle z^n, z^m\rangle = \int_{\mathbb{D}} z^n\overline{z^m}\,dA = \delta_{nm}\int_{\mathbb{D}} |z|^{2n}\,dA = \delta_{nm}\,\frac{\pi}{n+1}.$$

Hence

$$\phi_n(z) := \sqrt{\frac{n+1}{\pi}}\,z^n \qquad (n \ge 0)$$

is orthonormal, and analytic polynomials are dense in $L_a^2(\mathbb{D})$, so $\{\phi_n\}_{n\ge 0}$ is an ONB.
(b) Norm in terms of Taylor coefficients. If $f(z) = \sum_{k=0}^{\infty} a_k z^k$ in $\mathbb{D}$, then $f = \sum_{k\ge 0} c_k\phi_k$ with $c_k = a_k\sqrt{\pi/(k+1)}$, so

$$\|f\|_2^2 = \sum_{k=0}^{\infty} |c_k|^2 = \pi\sum_{k=0}^{\infty}\frac{|a_k|^2}{k+1}.$$

(c) Riesz representer of evaluation. For $w \in \mathbb{D}$, the evaluation functional $f \mapsto f(w)$ is bounded and

$$f(w) = \langle f, K_w\rangle, \qquad K_w(z) = \sum_{n=0}^{\infty}\phi_n(z)\overline{\phi_n(w)} = \sum_{n=0}^{\infty}\frac{n+1}{\pi}(z\overline{w})^n = \frac{1}{\pi(1-z\overline{w})^2}.$$

(24) The Dirichlet space on $\mathbb{D}$

Define the *Dirichlet space*

$$\mathcal{D} := \left\{ f \in \mathcal{O}(\mathbb{D}) : \int_{\mathbb{D}} |f'(z)|^2 \, dA(z) < \infty \right\}, \qquad \langle f, g \rangle_{\mathcal{D}} := f(0)\overline{g(0)} + \int_{\mathbb{D}} f'(z)\overline{g'(z)} \, dA(z).$$

(a) $\mathcal{D}$ is a Hilbert space. Define

$$T : \mathcal{D} \to \mathbb{C} \times L^2_a(\mathbb{D}), \qquad T(f) := (f(0), f').$$

With the product inner product on $\mathbb{C} \times L^2_a(\mathbb{D})$, T is an isometry since

$$\langle f, g \rangle_{\mathcal{D}} = \langle T(f), T(g) \rangle_{\mathbb{C} \times L^2_a}.$$

It is surjective: given (a, h) with $h \in L^2_a(\mathbb{D})$, define

$$f(z) := a + \int_0^z h(\zeta) \, d\zeta,$$

which is analytic on $\mathbb{D}$ (path-independence since h is analytic) and satisfies $f(0) = a$, $f' = h$. Thus $\mathcal{D} \cong \mathbb{C} \times L^2_a(\mathbb{D})$ and is complete.

(c) An orthonormal basis of $\mathcal{D}$. For $n \geq 1$,

$$\|z^n\|^2_{\mathcal{D}} = \int_{\mathbb{D}} |(z^n)'|^2 \, dA = \int_{\mathbb{D}} |nz^{n-1}|^2 \, dA = n^2 \cdot \frac{\pi}{n} = \pi n,$$

and $\|1\|^2_{\mathcal{D}} = |1|^2 = 1$. Also $\langle z^n, z^m \rangle_{\mathcal{D}} = 0$ for $n \neq m$. Hence

$$e_0(z) := 1, \qquad e_n(z) := \frac{z^n}{\sqrt{\pi n}} \ (n \geq 1)$$

is orthonormal, and polynomials are dense in $\mathcal{D}$, so $\{e_n\}_{n \geq 0}$ is an ONB.

(d) Norm in terms of Taylor coefficients. If $f(z) = \sum_{k=0}^{\infty} a_k z^k$, then expanding in the ONB gives

$$\|f\|^2_{\mathcal{D}} = |a_0|^2 + \pi \sum_{k=1}^{\infty} k \, |a_k|^2.$$

(b) Evaluation is bounded. For $w \in \mathbb{D}$ and $f = \sum_{n \geq 0} c_n e_n$,

$$|f(w)| = \left| \sum_{n \geq 0} c_n e_n(w) \right| \leq \left(\sum_{n \geq 0} |c_n|^2 \right)^{1/2} \left(\sum_{n \geq 0} |e_n(w)|^2 \right)^{1/2} = \|f\|_{\mathcal{D}} \left(\sum_{n \geq 0} |e_n(w)|^2 \right)^{1/2}.$$

But

$$\sum_{n \geq 0} |e_n(w)|^2 = 1 + \sum_{n=1}^{\infty} \frac{|w|^{2n}}{\pi n} = 1 - \frac{1}{\pi} \log\bigl(1 - |w|^2\bigr),$$

so

$$|f(w)| \leq \|f\|_{\mathcal{D}} \left(1 - \frac{1}{\pi} \log(1 - |w|^2) \right)^{1/2}.$$

(c)(e) Suppose $w \in \mathbb{D}$. Find an explicit formula for $T_w \in \mathcal{D}$ such that

$$f(w) = \langle f, T_w \rangle \qquad \forall f \in \mathcal{D}.$$

By the Riesz representation theorem, for each $w \in \mathbb{D}$ there exists a unique $T_w \in \mathcal{D}$ such that $f(w) = \langle f, T_w \rangle$ for all $f \in \mathcal{D}$. Write

$$T_w(z) = \sum_{k=0}^{\infty} c_k e_k(z).$$

For each basis vector e_n,

$$e_n(w) = L_w(e_n) = \langle e_n, T_w \rangle = \overline{c_n}$$

(using orthonormality and that the inner product is linear in the first variable), hence $c_n = \overline{e_n(w)}$, and therefore

$$T_w(z) = \sum_{k=0}^{\infty} \overline{e_k(w)}\, e_k(z).$$

Substituting the basis functions $e_0(z) = 1$ and $e_k(z) = \dfrac{z^k}{\sqrt{\pi k}}$ for $k \geq 1$ gives

$$T_w(z) = 1 + \sum_{k=1}^{\infty} \frac{\overline{w}^k}{\sqrt{\pi k}} \cdot \frac{z^k}{\sqrt{\pi k}} = 1 + \frac{1}{\pi} \sum_{k=1}^{\infty} \frac{(z\overline{w})^k}{k}.$$

Recognizing $\sum_{k=1}^{\infty} \dfrac{x^k}{k} = -\log(1-x)$ for $|x| < 1$, we obtain (for $|z|, |w| < 1$)

$$T_w(z) = 1 - \frac{1}{\pi} \log\bigl(1 - \overline{w}z\bigr) \in \mathcal{D},$$

and it satisfies $f(w) = \langle f, T_w \rangle$ for all $f \in \mathcal{D}$.

Chapter 9: Real and Complex Measures

Section 9A: Total Variation

(9.1) Definition (Real and complex measures). Suppose $(X, \mathcal{S})$ is a measurable space.

- A function $\nu : \mathcal{S} \to \mathbb{F}$ is *countably additive* if

  $$\nu\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \nu(E_k)$$

  for every disjoint sequence $E_1, E_2, \ldots$ of sets in $\mathcal{S}$.

- A *real measure* on $(X, \mathcal{S})$ is a countably additive function $\nu : \mathcal{S} \to \mathbb{R}$.

- A *complex measure* on $(X, \mathcal{S})$ is a countably additive function $\nu : \mathcal{S} \to \mathbb{C}$.

(9.3) Result (Absolute convergence for a disjoint union). Suppose ν is a complex measure on $(X, \mathcal{S})$. Then:

(i) $\nu(\varnothing) = 0$.

(ii) $\sum_{k=1}^{\infty} |\nu(E_k)| < \infty$ for every disjoint sequence $E_1, E_2, \ldots$ in $\mathcal{S}$.

(9.4) Result (Measure determined by an L^1-function). Suppose μ is a positive measure on $(X, \mathcal{S})$ and $h \in L^1(\mu)$. Define $\nu : \mathcal{S} \to \mathbb{F}$ by

$$\nu(E) = \int_E h \, d\mu.$$

Then ν is a real measure if $\mathbb{F} = \mathbb{R}$, and a complex measure if $\mathbb{F} = \mathbb{C}$.

Proof. If $E_1, E_2, \ldots$ are disjoint, then

$$\nu\left(\bigcup_{k=1}^{\infty} E_k\right) = \int \left(\sum_{k=1}^{\infty} \chi_{E_k}\right) h \, d\mu = \sum_{k=1}^{\infty} \int_{E_k} h \, d\mu = \sum_{k=1}^{\infty} \nu(E_k).$$

□

(9.6) Definition ($h \, d\mu$). Suppose μ is a (positive) measure on $(X, \mathcal{S})$ and $h \in L^1(\mu)$. Then $h \, d\mu$ is the real or complex measure on $(X, \mathcal{S})$ defined by

$$(h \, d\mu)(E) = \int_E h \, d\mu.$$

(9.7) Result (Properties of complex measures). Suppose ν is a complex measure on $(X, \mathcal{S})$. Then:

(i) $\nu(E \setminus D) = \nu(E) - \nu(D)$ for all $D, E \in \mathcal{S}$ with $D \subseteq E$.

(ii) $\nu(D \cup E) = \nu(D) + \nu(E) - \nu(D \cap E)$ for all $D, E \in \mathcal{S}$.

(iii) If $E_1 \subseteq E_2 \subseteq \cdots$ are in $\mathcal{S}$, then

$$\nu\left(\bigcup_{k=1}^{\infty} E_k\right) = \lim_{k \to \infty} \nu(E_k).$$

(iv) If $E_1 \supseteq E_2 \supseteq \cdots$ are in $\mathcal{S}$, then

$$\nu\left(\bigcap_{k=1}^{\infty} E_k\right) = \lim_{k \to \infty} \nu(E_k).$$

(9.8) Definition (Total variation measure). Suppose ν is a $\mathbb{C}$-measure on $(X, \mathcal{S})$. The *total variation measure* is the function $|\nu| : \mathcal{S} \to [0, \infty]$ defined by

$$|\nu|(E) = \sup \left\{ |\nu(E_1)| + \cdots + |\nu(E_n)| \ : \ n \in \mathbb{Z}^+, \ E_1, \ldots, E_n \text{ disjoint in } \mathcal{S}, \ \bigcup_{j=1}^{n} E_j \subseteq E \right\}.$$

(9.9) Result (Total variation measure of an $\mathbb{R}$-measure). Suppose ν is a real measure on $(X, \mathcal{S})$ and $E \in \mathcal{S}$. Then

$$|\nu|(E) = \sup \{ |\nu(A)| + |\nu(B)| \ : \ A, B \in \mathcal{S} \text{ disjoint and } A \cup B \subseteq E \}.$$

(9.10) Result (Total variation measure of $h\,d\mu$). Suppose μ is a (positive) measure on $(X, \mathcal{S})$, $h \in L^1(\mu)$, and $d\nu = h\,d\mu$. Then for all $E \in \mathcal{S}$,

$$|\nu|(E) = \int_E |h|\,d\mu.$$

(9.11) Result (Total variation measure is a measure). Suppose ν is a complex measure on $(X, \mathcal{S})$. Then the total variation $|\nu|$ is a (positive) measure on $(X, \mathcal{S})$.

Proof (sketch). Clearly $|\nu|(\varnothing) = 0$. Let $A_1, A_2, \ldots$ be disjoint sets in $\mathcal{S}$ and fix $m \in \mathbb{Z}^+$. For each $k \in \{1, \ldots, m\}$ choose disjoint sets $E_{k,1}, \ldots, E_{k,n_k} \in \mathcal{S}$ with $\bigcup_{i=1}^{n_k} E_{k,i} \subseteq A_k$. Then $\{E_{k,i}\}_{k,i}$ is disjoint and contained in $\bigcup_{k=1}^m A_k$, so

$$\sum_{k=1}^m \sum_{i=1}^{n_k} |\nu(E_{k,i})| \le |\nu|\left(\bigcup_{k=1}^m A_k\right).$$

Taking the supremum over all such choices gives

$$\sum_{k=1}^m |\nu|(A_k) \le |\nu|\left(\bigcup_{k=1}^m A_k\right).$$

Conversely, if $E_1, \ldots, E_n$ are disjoint with $\bigcup_{j=1}^n E_j \subseteq \bigcup_{k=1}^m A_k$, then

$$\sum_{j=1}^n |\nu(E_j)| = \sum_{j=1}^n \left|\sum_{k=1}^m \nu(E_j \cap A_k)\right| \le \sum_{k=1}^m \sum_{j=1}^n |\nu(E_j \cap A_k)| \le \sum_{k=1}^m |\nu|(A_k),$$

and taking the supremum over $E_1, \ldots, E_n$ yields

$$|\nu|\left(\bigcup_{k=1}^m A_k\right) \le \sum_{k=1}^m |\nu|(A_k).$$

Hence $|\nu|(\bigcup_{k=1}^m A_k) = \sum_{k=1}^m |\nu|(A_k)$ for every m, and letting $m \to \infty$ (using continuity from below) gives countable additivity. $\square$

(9.13) Definition (Addition and scalar multiplication of measures). Suppose μ, ν are $\mathbb{C}$-measures on $(X, \mathcal{S})$ and $\alpha \in \mathbb{C}$. Define the $\mathbb{C}$-measures $(\mu + \nu)$ and $(\alpha\nu)$ on $(X, \mathcal{S})$ by

$$(\mu + \nu)(E) = \mu(E) + \nu(E), \qquad (\alpha\nu)(E) = \alpha\,\nu(E).$$

(9.14) Definition ($\mathcal{M}_{\mathbb{F}}(X, \mathcal{S})$). Suppose $(X, \mathcal{S})$ is a measurable space. Then $\mathcal{M}_{\mathbb{F}}(X, \mathcal{S})$ denotes the vector space of $\mathbb{F}$-measures on $(X, \mathcal{S})$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$.

(9.15) Definition (Total variation norm of a $\mathbb{C}$-measure). Suppose ν is a $\mathbb{C}$-measure on $(X, \mathcal{S})$. The *total variation norm* of ν, denoted $\|\nu\|$, is defined by

$$\|\nu\| := |\nu|(X).$$

(9.17) Result (Total variation norm is finite). Suppose $(X, \mathcal{S})$ is a measurable space and $\nu \in \mathcal{M}_{\mathbb{C}}(X, \mathcal{S})$. Then

$$\|\nu\| < \infty.$$

(9.18) Result (Measures form a Banach space). For a measurable space $(X, \mathcal{S})$, the space $\mathcal{M}_{\mathbb{F}}(X, \mathcal{S})$ is a Banach space with the total variation norm $\|\cdot\|$.

Proof (standard completeness argument). Let $\nu_1, \nu_2, \ldots$ be a Cauchy sequence in $\mathcal{M}_{\mathbb{F}}(X, \mathcal{S})$. For every $E \in \mathcal{S}$,

$$|\nu_j(E) - \nu_k(E)| = |(\nu_j - \nu_k)(E)| \le |\nu_j - \nu_k|(E) \le \|\nu_j - \nu_k\|.$$

Hence $(\nu_n(E))_{n\ge 1}$ is Cauchy in $\mathbb{F}$, so it converges. Define $\nu : \mathcal{S} \to \mathbb{F}$ by

$$\nu(E) := \lim_{n\to\infty} \nu_n(E).$$

We claim $\nu \in \mathcal{M}_{\mathbb{F}}(X, \mathcal{S})$ and $\nu_n \to \nu$ in $\|\cdot\|$.

First, ν is countably additive. Let $E_1, E_2, \ldots$ be disjoint in $\mathcal{S}$. Fix $\varepsilon > 0$ and choose m such that $\|\nu_j - \nu_m\| \le \varepsilon$ for all $j \ge m$. Since ν_m is a (real/complex) measure, the series $\sum_{k=1}^{\infty} |\nu_m(E_k)|$ converges, so pick N with $\sum_{k=N}^{\infty} |\nu_m(E_k)| < \varepsilon$. For $j \ge m$,

$$\begin{aligned}\sum_{k=N}^{\infty} |\nu_j(E_k)| &\le \sum_{k=N}^{\infty} |\nu_j(E_k) - \nu_m(E_k)| + \sum_{k=N}^{\infty} |\nu_m(E_k)| \\ &\le \sum_{k=N}^{\infty} |\nu_j - \nu_m|(E_k) + \varepsilon = |\nu_j - \nu_m|\left(\bigcup_{k=N}^{\infty} E_k\right) + \varepsilon \\ &\le \|\nu_j - \nu_m\| + \varepsilon \le 2\varepsilon.\end{aligned}$$

Letting $j \to \infty$ gives $\sum_{k=N}^{\infty} |\nu(E_k)| \le 2\varepsilon$, so $\sum_{k=1}^{\infty} \nu(E_k)$ converges absolutely. Moreover, since $\nu_m(\bigcup_{k=1}^{\infty} E_k) = \sum_{k=1}^{\infty} \nu_m(E_k)$ and $\nu_j \to \nu$ pointwise on sets,

$$\nu\left(\bigcup_{k=1}^{\infty} E_k\right) = \lim_{j\to\infty} \nu_j\left(\bigcup_{k=1}^{\infty} E_k\right) = \lim_{j\to\infty} \sum_{k=1}^{\infty} \nu_j(E_k) = \sum_{k=1}^{\infty} \nu(E_k),$$

where the exchange of limit and sum is justified by the uniform tail bound above. Hence ν is a measure.

Finally, $\nu_n \to \nu$ in total variation norm. Fix $\varepsilon > 0$ and choose m such that $\|\nu_j - \nu_m\| \le \varepsilon$ for all $j \ge m$. For any disjoint sequence $(E_k) \subseteq \mathcal{S}$,

$$\sum_{k=1}^{\infty} |(\nu - \nu_m)(E_k)| = \lim_{j\to\infty} \sum_{k=1}^{\infty} |(\nu_j - \nu_m)(E_k)| \le \limsup_{j\to\infty} \|\nu_j - \nu_m\| \le \varepsilon.$$

Taking the supremum over all such disjoint sequences (as in the definition of total variation) yields $\|\nu - \nu_m\| \le \varepsilon$. Thus $\nu_n \to \nu$ in norm, proving completeness. $\square$

Exercises 9A

4) Suppose ν is a $\mathbb{C}$-measure on $(X, \mathcal{S})$. Prove that if $E \in \mathcal{S}$, then

$$|\nu|(E) = \sup\left\{\sum_{k=1}^{\infty} |\nu(E_k)| \;:\; E_1, E_2, \ldots \text{ disjoint in } \mathcal{S},\ E = \bigcup_{k=1}^{\infty} E_k\right\}.$$

Proof. Recall

$$|\nu|(E) = \sup\left\{\sum_{j=1}^{n} |\nu(F_j)| \ : \ n \in \mathbb{Z}^+, \ F_1, \ldots, F_n \text{ disjoint in } \mathcal{S}, \ \bigcup_{j=1}^{n} F_j \subseteq E\right\}.$$

First note the supremum is unchanged if we require $\bigcup_{j=1}^n F_j = E$: if $F_1, \ldots, F_n$ are disjoint with $\bigcup_{j=1}^n F_j \subseteq E$, set $F_{n+1} := E \setminus \bigcup_{j=1}^n F_j$. Then $F_1, \ldots, F_{n+1}$ are disjoint, $\bigcup_{j=1}^{n+1} F_j = E$, and

$$\sum_{j=1}^{n+1} |\nu(F_j)| \geq \sum_{j=1}^{n} |\nu(F_j)|.$$

Hence

$$|\nu|(E) = \sup\left\{\sum_{j=1}^{n} |\nu(F_j)| \ : \ F_1, \ldots, F_n \text{ disjoint}, \ \bigcup_{j=1}^{n} F_j = E\right\}. \qquad (*)$$

Define

$$T(E) := \sup\left\{\sum_{k=1}^{\infty} |\nu(E_k)| \ : \ (E_k)_{k \geq 1} \text{ disjoint in } \mathcal{S}, \ E = \bigcup_{k=1}^{\infty} E_k\right\}.$$

(i) $T(E) \leq |\nu|(E)$. If $E = \bigcup_{k=1}^\infty E_k$ with E_k disjoint, then $|\nu(E_k)| \leq |\nu|(E_k)$ for each k, and since $|\nu|$ is a measure,

$$\sum_{k=1}^{\infty} |\nu(E_k)| \leq \sum_{k=1}^{\infty} |\nu|(E_k) = |\nu|\left(\bigcup_{k=1}^{\infty} E_k\right) = |\nu|(E).$$

Taking the supremum over all such sequences gives $T(E) \leq |\nu|(E)$.

(ii) $T(E) \geq |\nu|(E)$. Let $F_1, \ldots, F_n$ be a disjoint partition of E. Define a disjoint sequence (E_k) by $E_k = F_k$ for $1 \leq k \leq n$ and $E_k = \varnothing$ for $k > n$. Then $\bigcup_{k=1}^\infty E_k = E$ and

$$\sum_{k=1}^{\infty} |\nu(E_k)| = \sum_{k=1}^{n} |\nu(F_k)|.$$

Thus every sum appearing in $(*)$ also appears among the candidates for $T(E)$, so $T(E) \geq |\nu|(E)$.

Combining (i) and (ii) yields $T(E) = |\nu|(E)$. ☐

5) Let $(X, \mathcal{S}, \mu)$ be a measure space and let $h \geq 0$ with $h \in L^1(\mu)$. Define $\nu(E) := \int_E h \, d\mu$ for $E \in \mathcal{S}$. Then ν is a (positive) measure, and for every $\mathcal{S}$-measurable $f \geq 0$,

$$\int f \, d\nu = \int f \, h \, d\mu.$$

Proof. For $f = \mathbf{1}_A$ with $A \in \mathcal{S}$,

$$\int \mathbf{1}_A \, d\nu = \nu(A) = \int_A h \, d\mu = \int \mathbf{1}_A \, h \, d\mu.$$

If $f = \sum_{k=1}^n a_k \mathbf{1}_{E_k}$ is a nonnegative simple function with $a_k \geq 0$ and pairwise disjoint $E_k \in \mathcal{S}$, then by linearity,

$$\int f \, d\nu = \sum_{k=1}^{n} a_k \nu(E_k) = \sum_{k=1}^{n} a_k \int_{E_k} h \, d\mu = \int \left(\sum_{k=1}^{n} a_k \mathbf{1}_{E_k}\right) h \, d\mu = \int f \, h \, d\mu.$$

For general measurable $f \geq 0$, choose simple functions $f_n \geq 0$ with $f_n \uparrow f$ pointwise. Then $f_n h \uparrow fh$, so by the monotone convergence theorem,

$$\int f\,d\nu = \lim_{n\to\infty} \int f_n\,d\nu = \lim_{n\to\infty} \int f_n h\,d\mu = \int fh\,d\mu.$$

$\square$

Section 9B: Decomposition Theorems

(9.23) Result (Hahn Decomposition Theorem). Suppose ν is an $\mathbb{R}$-measure on $(X, \mathcal{S})$. Then there exist sets $A, B \in \mathcal{S}$ such that:

(i) $A \cup B = X$ and $A \cap B = \varnothing$,

(ii) $\nu(E) \geq 0$ for all $E \in \mathcal{S}$ with $E \subseteq A$,

(iii) $\nu(E) \leq 0$ for all $E \in \mathcal{S}$ with $E \subseteq B$.

In other words, a real measure on $(X, \mathcal{S})$ decomposes X into two disjoint measurable sets such that every measurable subset of one has nonnegative measure and every measurable subset of the other has nonpositive measure.

(9.28) Definition (Singular measures). Suppose μ and ν are complex or positive measures on $(X, \mathcal{S})$. Then μ and ν are called *singular with respect to each other* (denoted $\nu \perp \mu$) if there exist sets $A, B \in \mathcal{S}$ such that:

(i) $A \cup B = X$ and $A \cap B = \varnothing$,

(ii) $\nu(E) = \nu(E \cap A)$ and $\mu(E) = \mu(E \cap B)$ for all $E \in \mathcal{S}$.

In other words, two complex (or positive) measures are singular if the two measures "live on different sets."

Example. Suppose λ is Lebesgue measure on the σ-algebra $\mathcal{B}$ of Borel subsets of $\mathbb{R}$. Let $r_1, r_2, \dots \in \mathbb{Q}$, and let $w_1, w_2, \dots$ be a bounded sequence of complex numbers. Define a complex measure ν on $(\mathbb{R}, \mathcal{B})$ by

$$\nu(E) := \sum_{r_n \in E} \frac{w_n}{2^n}, \qquad E \in \mathcal{B}.$$

Then $\nu \perp \lambda$ because ν is supported on $\mathbb{Q}$ and λ is supported on $\mathbb{R} \setminus \mathbb{Q}$.

(9.30) Result (Jordan Decomposition Theorem). Every $\mathbb{R}$-measure is the difference of two finite (positive) measures that are singular with respect to each other. More precisely: if ν is an $\mathbb{R}$-measure on $(X, \mathcal{S})$, then there exist unique finite (positive) measures ν^+, ν^- on $(X, \mathcal{S})$ such that

$$\nu = \nu^+ - \nu^-, \qquad \nu^+ \perp \nu^-, \qquad |\nu| = \nu^+ + \nu^-.$$

(9.32) Definition (Absolutely continuous). Suppose ν is a $\mathbb{C}$-measure on $(X, \mathcal{S})$ and μ is a (positive) measure on $(X, \mathcal{S})$. Then ν is called *absolutely continuous with respect to* μ (denoted $\nu \ll \mu$) if

$$\nu(E) = 0 \quad \forall E \in \mathcal{S} \text{ with } \mu(E) = 0.$$

Examples:

(i) If μ is a positive measure and $h \in L^1(\mu)$, then $h\, d\mu \ll \mu$.

(ii) If ν is an $\mathbb{R}$-measure, then $\nu^+ \ll |\nu|$ and $\nu^- \ll |\nu|$.

(iii) If ν is a $\mathbb{C}$-measure, then $\nu \ll |\nu|$, $\Re(\nu) \ll |\nu|$, and $\Im(\nu) \ll |\nu|$.

(iv) Every measure on $(X, \mathcal{S})$ is absolutely continuous with respect to counting measure on $(X, \mathcal{S})$.

(9.35) Result (Lebesgue Decomposition Theorem). Suppose μ is a (positive) measure on $(X, \mathcal{S})$.

(i) Every $\mathbb{C}$-measure on $(X, \mathcal{S})$ is the sum of a $\mathbb{C}$-measure absolutely continuous with respect to μ and a $\mathbb{C}$-measure singular with respect to μ.

More precisely: if ν is a $\mathbb{C}$-measure on $(X, \mathcal{S})$, then there exist unique $\mathbb{C}$-measures ν_a, ν_s on $(X, \mathcal{S})$ such that

$$\nu = \nu_a + \nu_s, \qquad \nu_a \ll \mu, \qquad \nu_s \perp \mu.$$

In other words, a (positive) measure μ determines a decomposition of each complex measure into the sum of the two extreme types (absolutely continuous and singular).

(9.36) Result (Radon–Nikodym Theorem). Suppose μ is a (positive) σ-finite measure on $(X, \mathcal{S})$ and suppose ν is a $\mathbb{C}$-measure on $(X, \mathcal{S})$ such that $\nu \ll \mu$. Then there exists $h \in L^1(\mu)$ such that

$$d\nu = h\, d\mu.$$

Equivalently: if μ is σ-finite, then every $\mathbb{C}$-measure absolutely continuous with respect to μ is of the form $h\, d\mu$ for some $h \in L^1(\mu)$.

(9.41) Result (Polar decomposition: $d\nu = h\, d|\nu|$). If ν is a $\mathbb{C}$-measure on $(X, \mathcal{S})$, then $d\nu = h\, d|\nu|$ for some measurable h with $|h(x)| = 1$.

(i) If ν is an $\mathbb{R}$-measure on $(X, \mathcal{S})$, then there exists an $\mathcal{S}$-measurable function $h : X \to \{-1, 1\}$ such that $d\nu = h\, d|\nu|$.

(ii) If ν is a $\mathbb{C}$-measure on $(X, \mathcal{S})$, then there exists an $\mathcal{S}$-measurable function $h : X \to \{z \in \mathbb{C} : |z| = 1\}$ such that $d\nu = h\, d|\nu|$.

(9.42) Result (Dual space of $L^p(\mu)$ is $L^{p'}(\mu)$). Suppose μ is a (positive) measure and $1 \le p < \infty$ (with the additional hypothesis that μ is σ-finite if $p = 1$). Let p' be the conjugate exponent, i.e. $\frac{1}{p} + \frac{1}{p'} = 1$. For $h \in L^{p'}(\mu)$ define $\varphi_h : L^p(\mu) \to \mathbb{F}$ by

$$\varphi_h(f) = \int f h\, d\mu.$$

Then $h \mapsto \varphi_h$ is an injective linear map from $L^{p'}(\mu)$ onto $(L^p(\mu))^*$. Furthermore,

$$\|\varphi_h\| = \|h\|_{p'} \qquad \text{for all } h \in L^{p'}(\mu).$$

Exercises 9B

2) Suppose μ is a (positive) measure and $g, h \in L^2(\mu)$. Prove that

$$g\,d\mu \perp h\,d\mu \quad\Longleftrightarrow\quad g(x)h(x) = 0 \text{ for } \mu\text{-a.e. } x \in X.$$

Proof. ($\Rightarrow$) If $g\,d\mu \perp h\,d\mu$, then there exists $A \in \mathcal{S}$ such that

$$|g\,d\mu|(A) = 0, \qquad |h\,d\mu|(X \setminus A) = 0.$$

But $|g\,d\mu|(A) = \int_A |g|\,d\mu$, so $g = 0$ μ-a.e. on A; similarly $h = 0$ μ-a.e. on $X \setminus A$. Hence $g(x)h(x) = 0$ for μ-a.e. $x \in X$.

($\Leftarrow$) Assume $g(x)h(x) = 0$ for μ-a.e. x. Set $A := \{x \in X : \ g(x) = 0\}$. Then $g = 0$ on A and, on $X \setminus A$, we have $g \neq 0$ so $h = 0$ μ-a.e. on $X \setminus A$ (by the assumption $gh = 0$ a.e.). Therefore,

$$|g\,d\mu|(A) = \int_A |g|\,d\mu = 0, \qquad |h\,d\mu|(X \setminus A) = \int_{X\setminus A} |h|\,d\mu = 0,$$

which shows $g\,d\mu \perp h\,d\mu$. □

Exercises 9B (continued)

7) Use the Cantor set to prove that there exists a (positive) measure ν on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that

$$\nu \perp \lambda, \qquad \nu(\mathbb{R}) \neq 0, \qquad \text{but } \nu(\{x\}) = 0 \ \ \forall x \in \mathbb{R},$$

where λ denotes Lebesgue measure and $\mathcal{B}(\mathbb{R})$ the Borel σ-algebra.

Proof. Let $C \subset [0,1]$ be the Cantor set. Recall $\lambda(C) = 0$ and $C = \bigcap_{n\geq 0} C_n$, where $C_0 = [0,1]$ and C_n is obtained from C_{n-1} by removing the middle third of each interval. At stage n, the set C_n is a union of 2^n pairwise disjoint closed intervals, each of length 3^{-n}; write these as

$$C_n = \bigcup_{k=1}^{2^n} I_k^{(n)}.$$

(i) Define a premeasure on the Cantor-interval algebra. Let $\mathcal{A}$ be the algebra of finite disjoint unions of intervals of the form $I_k^{(n)}$. Define $\nu_0 : \mathcal{A} \to [0,\infty)$ by declaring

$$\nu_0\Big(I_k^{(n)}\Big) := 2^{-n}, \qquad \nu_0\left(\bigcup_{j=1}^m J_j\right) := \sum_{j=1}^m \nu_0(J_j) \ \text{ for disjoint } J_j \in \{I_k^{(n)}\}.$$

This is well-defined: if $m \geq n$, then each $I_k^{(n)}$ is the disjoint union of exactly 2^{m-n} intervals of level m, so any refinement preserves total mass since

$$2^{m-n} \cdot 2^{-m} = 2^{-n}.$$

Hence ν_0 is a finitely additive premeasure on $\mathcal{A}$, and

$$\nu_0([0,1]) = \nu_0(C_n) = \sum_{k=1}^{2^n} \nu_0(I_k^{(n)}) = 2^n \cdot 2^{-n} = 1.$$

By Carathéodory's extension theorem, ν_0 extends to a measure on $\sigma(\mathcal{A})$. Since the triadic intervals generate the Borel σ-algebra on $[0,1]$, we may view this extension as a Borel measure on $([0,1], \mathcal{B}([0,1]))$, still denoted ν.

Extend ν to $\mathbb{R}$ by

$$\nu(E) := \nu(E \cap [0,1]) \qquad \forall E \in \mathcal{B}(\mathbb{R}).$$

(ii) $\nu(\mathbb{R}) \neq 0$ and $\nu \perp \lambda$. By construction,

$$\nu(\mathbb{R}) = \nu([0,1]) = 1 \neq 0.$$

Also, for each n,

$$[0,1] = C_n \,\dot{\cup}\, ([0,1] \setminus C_n), \qquad \nu(C_n) = 1 \implies \nu([0,1] \setminus C_n) = 0.$$

The sets $[0,1] \setminus C_n$ increase to $[0,1] \setminus C$, so by continuity from below,

$$\nu([0,1] \setminus C) = \lim_{n\to\infty} \nu([0,1] \setminus C_n) = 0.$$

Hence $\nu(\mathbb{R} \setminus C) = 0$ (since ν is supported on $[0,1]$), while $\lambda(C) = 0$. Therefore $\nu \perp \lambda$.

(iii) $\nu(\{x\}) = 0$ for all $x \in \mathbb{R}$. If $x \notin [0,1]$, then $\nu(\{x\}) = \nu(\varnothing) = 0$. If $x \in [0,1] \setminus C$, then $\nu(\{x\}) \leq \nu([0,1] \setminus C) = 0$. If $x \in C$, then for each n there is a unique interval $I^{(n)}(x) \in \{I_k^{(n)}\}_{k=1}^{2^n}$ with $x \in I^{(n)}(x)$. These form a nested sequence

$$I^{(1)}(x) \supset I^{(2)}(x) \supset \cdots, \qquad \bigcap_{n\geq 1} I^{(n)}(x) = \{x\}.$$

Since $\nu(I^{(n)}(x)) = 2^{-n}$ and ν is finite on $[0,1]$, continuity from above gives

$$\nu(\{x\}) = \nu\left(\bigcap_{n\geq 1} I^{(n)}(x)\right) = \lim_{n\to\infty} \nu(I^{(n)}(x)) = \lim_{n\to\infty} 2^{-n} = 0.$$

This completes the construction. $\square$

Chapter 10: Linear Maps on Hilbert Spaces

10A: Adjoints and Invertibility

(10.1) Definition (Adjoint, T^*). Suppose V, W are Hilbert spaces and $T : V \to W$ is a bounded linear map. The *adjoint* of T is the map $T^* : W \to V$ such that

$$\langle Tf, g\rangle_W = \langle f, T^*g\rangle_V \qquad \forall f \in V,\ \forall g \in W.$$

(Here the inner product on the left is in W and on the right is in V.) Moreover, one has the estimate

$$\|T^*g\| \leq \|T\|\,\|g\| \qquad \forall g \in W.$$

Example (Multiplication operators). Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $h \in L^\infty(\mu)$. Define $M_h : L^2(\mu) \to L^2(\mu)$ by $M_h f = hf$. Then M_h is bounded and $\|M_h\| \leq \|h\|_\infty$. Also,

$$\langle M_h f, g\rangle = \int hf\,\overline{g}\,d\mu = \int f\,\overline{\overline{h}g}\,d\mu = \langle f, M_{\overline{h}}g\rangle,$$

so $M_h^* = M_{\overline{h}}$.

Example (Integral operators). Let $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{T}, \nu)$ be σ-finite measure spaces and let $K \in L^2(\mu \times \nu)$. Define $I_K : L^2(\nu) \to L^2(\mu)$ by

$$(I_K f)(x) := \int_Y K(x,y) f(y)\, d\nu(y) \quad \text{for a.e. } x.$$

By Cauchy–Schwarz,

$$|I_K f(x)|^2 \le \left(\int_Y |K(x,y)|^2 \, d\nu(y) \right) \|f\|_2^2.$$

Integrating in x and using Tonelli yields

$$\|I_K f\|_2^2 \le \|K\|_{L^2(\mu\times\nu)}^2 \, \|f\|_2^2,$$

so I_K is bounded and $\|I_K\| \le \|K\|_2$. If $K^* : Y \times X \to \mathbb{F}$ is defined by $K^*(y,x) := \overline{K(x,y)}$, then

$$\langle I_K f, g\rangle_{L^2(\mu)} = \langle f, I_{K^*} g\rangle_{L^2(\nu)},$$

hence $(I_K)^* = I_{K^*}$. (Think of I_K as an "infinite matrix" acting by kernel multiplication.)

Example (Matrices). Take $X = \{1, \dots, m\}$, $Y = \{1, \dots, n\}$ with counting measures, and let $K = (K(i,j))$ be an $m \times n$ matrix. Then $L^2(\nu) \cong \mathbb{F}^n$, $L^2(\mu) \cong \mathbb{F}^m$, and

$$(I_K f)(i) = \sum_{j=1}^{n} K(i,j) f(j).$$

Here $K^*(j,i) = \overline{K(i,j)}$ and $(I_K)^* = I_{K^*}$ (multiplication by the conjugate-transpose). A norm estimate is

$$\|I_K\| \le \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |K(i,j)|^2 \right)^{1/2}.$$

(10.11) Result (T^* is bounded). Suppose V, W are Hilbert spaces and $T \in \mathcal{B}(V,W)$ (the Banach space of bounded linear maps $V \to W$). Then $T^* \in \mathcal{B}(W,V)$, $(T^*)^* = T$, and

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

In other words, if $T : V \to W$, then $T^* : W \to V$.

(10.12) Result (Properties of the adjoint). Suppose V, W, U are Hilbert spaces. Then:

(i) $(S+T)^* = S^* + T^*$ for all $S, T \in \mathcal{B}(V,W)$.

(ii) $(\alpha T)^* = \overline{\alpha}\, T^*$ for all $\alpha \in \mathbb{F}$ and $T \in \mathcal{B}(V,W)$.

(iii) $I^* = I$ where I is the identity operator on V.

(iv) $(S \circ T)^* = T^* \circ S^*$ for $T \in \mathcal{B}(V,W)$ and $S \in \mathcal{B}(W,U)$.

(10.13) Result (Null space and range of T^*). Suppose V, W are Hilbert spaces and $T \in \mathcal{B}(V,W)$. Then

$$\ker(T^*) = (\operatorname{range} T)^{\perp}, \qquad \operatorname{range}(T^*) = (\ker T)^{\perp},$$

and equivalently,

$$\ker(T) = (\operatorname{range} T^*)^{\perp}, \qquad \operatorname{range}(T) = (\ker T^*)^{\perp}.$$

Continued – 10A

(10.14) Result (Condition for dense range). Suppose V, W are Hilbert spaces and $T \in B(V, W)$. Then T has dense range if and only if T^* is injective (one-to-one). Equivalently,

$$\overline{\operatorname{Ran}(T)} = (\ker T^*)^\perp \quad \Longrightarrow \quad \overline{\operatorname{Ran}(T)} = W \iff \ker(T^*) = \{0\}.$$

Advantage: to check whether a bounded linear map between Hilbert spaces has dense range, it suffices to check whether $T^*g = 0$ implies $g = 0$.

(10.17) Definition (Operator; $B(V)$).

- An *operator* is a linear map from a vector space to itself.
- If V is a normed vector space, $B(V)$ denotes the normed vector space of bounded linear operators $V \to V$.
- More generally, $B(U, V)$ denotes the bounded linear maps $U \to V$ (so $B(V) = B(V, V)$).

(10.18) Definition (Invertible; T^{-1}). An operator $T : V \to V$ is *invertible* if T is bijective. Equivalently, T is invertible iff there exists an operator $T^{-1} : V \to V$ such that

$$T^{-1}T = TT^{-1} = I.$$

(10.19) Result (Inverse of the adjoint = adjoint of the inverse). If T is a bounded operator on a Hilbert space, then

$$T \text{ invertible} \iff T^* \text{ invertible}, \qquad (T^*)^{-1} = (T^{-1})^*.$$

(10.20) Result (Norm of a composition). Suppose U, V, W are normed vector spaces, $T \in B(U, V)$ and $S \in B(V, W)$. Then

$$\|ST\| \leq \|S\| \, \|T\|.$$

(10.21) Definition (T^k). Let T be an operator on a vector space V.

- For $k \in \mathbb{Z}_+$, define $T^k := \underbrace{T \circ T \circ \cdots \circ T}_{k \text{ times}}$.
- Define $T^0 := I : V \to V$.
- $T^jT^k = T^{j+k}$ and $(T^j)^k = T^{jk}$.
- If V is normed and $T \in B(V)$, then $\|T^k\| \leq \|T\|^k$.

(10.22) Result (Neumann series; ball around I consists of invertibles). If T is a bounded operator on a Banach space and $\|T\| < 1$, then $I - T$ is invertible and

$$(I - T)^{-1} = \sum_{k=0}^{\infty} T^k$$

(where the series converges in operator norm).

Proof sketch. Since $\sum_{k=0}^{\infty} \|T^k\| \le \sum_{k=0}^{\infty} \|T\|^k = \frac{1}{1-\|T\|} < \infty$, the operator series converges in $B(V)$. Moreover,

$$(I - T) \sum_{k=0}^{n} T^k = I - T^{n+1} \xrightarrow[n\to\infty]{} I,$$

so $(I - T) \left(\sum_{k=0}^{\infty} T^k\right) = I$ (and similarly on the other side). $\square$

(10.25) Result (Invertible operators form an open set). Suppose V is a Banach space. Then

$$\mathrm{Inv}(V) := \{T \in B(V) : T \text{ invertible}\}$$

is an open subset of $B(V)$.

(10.26) Definition (Left invertible; right invertible). Suppose T is a bounded operator on a Banach space V.

- T is *left invertible* if there exists $S \in B(V)$ such that $ST = I$.
- T is *right invertible* if there exists $S \in B(V)$ such that $TS = I$.

In finite-dimensional linear algebra, left and right invertibility are equivalent; this can fail on infinite-dimensional spaces.

Equivalent conditions for left invertibility (Hilbert space case). Suppose V is a Hilbert space and $T \in B(V)$. The following are equivalent:

- T is left invertible.
- $\exists\, a > 0$ such that $\|f\| \le a\|Tf\|$ for all $f \in V$ (equivalently, $\|Tf\| \ge c\|f\|$ for some $c > 0$).
- T is injective and $\mathrm{Ran}(T)$ is closed.
- T^*T is invertible.

Example (closed range requirement). Define $T : \ell^2 \to \ell^2$ by

$$T(a_1, a_2, a_3, \dots) = \left(a_1, \frac{a_2}{2}, \frac{a_3}{3}, \dots\right).$$

Then T is bounded and injective, but it is *not* left invertible (its range is not closed). Indeed, if $S \in B(\ell^2)$ satisfied $ST = I$, then for the standard basis vectors e_n we have

$$e_n = n\,T(e_n) \quad \Rightarrow \quad S(e_n) = n\,S(T(e_n)) = n\,e_n,$$

so $\|S(e_n)\| = n \to \infty$, contradicting boundedness of S.

Equivalent conditions for right invertibility (Hilbert space case). Suppose V is a Hilbert space and $T \in B(V)$. The following are equivalent:

- T is right invertible.
- T is surjective.
- TT^* is invertible.

Exercises – 10A

3) Suppose V, W are Hilbert spaces and $g \in V$, $h \in W$. Define $T \in B(V, W)$ by

$$Tf = \langle f, g \rangle\, h.$$

Find a formula for T^*.
For $f \in V$ and $k \in W$,

$$\langle Tf, k \rangle_W = \langle \langle f, g \rangle h,\ k \rangle_W = \langle f, g \rangle \langle h, k \rangle_W.$$

To have $\langle Tf, k \rangle_W = \langle f, T^* k \rangle_V$ for all f, take $T^* k = \lambda(k)\, g$ and solve:

$$\langle f, \lambda(k) g \rangle_V = \overline{\lambda(k)}\, \langle f, g \rangle = \langle f, g \rangle \langle h, k \rangle_W \quad \Rightarrow \quad \overline{\lambda(k)} = \langle h, k \rangle_W.$$

Hence $\lambda(k) = \overline{\langle h, k \rangle_W} = \langle k, h \rangle_W$, so

$$\boxed{T^* k = \langle k, h \rangle_W\, g, \qquad k \in W.}$$

5) Prove or give a counterexample: If V is a Hilbert space and $T : V \to V$ is bounded with $\dim(\ker T) < \infty$, then $\dim(\ker T^*) < \infty$.
Counterexample. Let $V = \ell^2(\mathbb{Z}_+)$ and define $T : \ell^2 \to \ell^2$ by

$$T(x_1, x_2, x_3, \dots) = (0, x_1, 0, x_2, 0, x_3, \dots),$$

i.e. $(Tx)_{2n} = x_n$ and $(Tx)_{2n-1} = 0$. Then $\|Tx\|^2 = \sum_{n \geq 1} |x_n|^2 = \|x\|^2$, so T is an isometry and $\ker(T) = \{0\}$ (hence $\dim \ker(T) = 0 < \infty$). But

$$\operatorname{Ran}(T) = \{y \in \ell^2 : y_{2n-1} = 0\ \forall n\} = \overline{\operatorname{span}}\{e_2, e_4, e_6, \dots\},$$

so

$$\ker(T^*) = (\operatorname{Ran}(T))^\perp = \overline{\operatorname{span}}\{e_1, e_3, e_5, \dots\},$$

which is infinite-dimensional. Therefore $\dim(\ker T^*) = \infty$.

7) Suppose V is a Hilbert space and $\operatorname{Inv}(V)$ is the set of invertible bounded operators on V. Show that $T \mapsto T^{-1}$ is continuous from $\operatorname{Inv}(V)$ to $\operatorname{Inv}(V)$.
Let $T \in \operatorname{Inv}(V)$ and let $S \in B(V)$ be close to T. Write

$$S = T + (S - T) = T\big(I + T^{-1}(S - T)\big).$$

Set $A := T^{-1}(S - T)$. If $\|A\| < 1$, then $I + A$ is invertible and

$$(I + A)^{-1} = \sum_{n=0}^{\infty} (-A)^n, \qquad \|(I + A)^{-1}\| \leq \frac{1}{1 - \|A\|}.$$

Hence $S^{-1} = (I + A)^{-1} T^{-1}$ and

$$\|S^{-1} - T^{-1}\| = \|(I + A)^{-1} T^{-1} - T^{-1}\| = \big\|\big((I + A)^{-1} - I\big) T^{-1}\big\| \leq \|(I + A)^{-1} - I\|\, \|T^{-1}\|.$$

Moreover, using $(I + A)^{-1} - I = -(I + A)^{-1} A$,

$$\|S^{-1} - T^{-1}\| \leq \|(I + A)^{-1}\|\, \|A\|\, \|T^{-1}\| \leq \frac{\|T^{-1}\|\, \|A\|}{1 - \|A\|} \leq \frac{\|T^{-1}\|^2\, \|S - T\|}{1 - \|T^{-1}\|\, \|S - T\|}.$$

Thus, given $\varepsilon > 0$, choosing $\|S - T\|$ sufficiently small forces $\|S^{-1} - T^{-1}\| < \varepsilon$. So inversion is continuous on $\operatorname{Inv}(V)$.

Section 10B – Spectrum

(10.32) Definition (Spectrum; sp(T); eigenvalue). Suppose T is a bounded operator on a Banach space V over $\mathbb{F}$.

- A number $\alpha \in \mathbb{F}$ is an *eigenvalue* of T if $T - \alpha I$ is not injective.
- A nonzero vector $f \in V$ is an *eigenvector* corresponding to α if $Tf = \alpha f$.
- The *spectrum* of T, denoted sp(T), is

$$\operatorname{sp}(T) := \{\alpha \in \mathbb{F} : T - \alpha I \text{ is not invertible}\}.$$

Examples (eigenvalues and spectrum).

(i) Let $(b_1, b_2, \dots)$ be a bounded sequence in $\mathbb{F}$ and define $T : \ell^2 \to \ell^2$ by

$$T(a_1, a_2, \dots) = (a_1 b_1, a_2 b_2, \dots).$$

Then the eigenvalues of T are $\{b_k : k \in \mathbb{Z}_+\}$ and

$$\operatorname{sp}(T) = \overline{\{b_k : k \in \mathbb{Z}_+\}}.$$

(ii) Suppose $h \in L^\infty(\mathbb{R})$ and define $M_h : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ by $M_h f = fh$. Then $\alpha \in \mathbb{F}$ is an eigenvalue of M_h iff

$$\big|\{t \in \mathbb{R} : h(t) = \alpha\}\big| > 0,$$

and

$$\alpha \in \operatorname{sp}(M_h) \iff \forall \varepsilon > 0,\ \big|\{t \in \mathbb{R} : |h(t) - \alpha| < \varepsilon\}\big| > 0.$$

(10.34) Result ($T - \alpha I$ invertible for $|\alpha|$ large). Suppose T is a bounded operator on a Banach space. Then:

(i) $\operatorname{sp}(T) \subseteq \{\alpha \in \mathbb{F} : |\alpha| \le \|T\|\}$.

(ii) $T - \alpha I$ is invertible for all α with $|\alpha| > \|T\|$.

(iii) $\displaystyle\lim_{|\alpha| \to \infty} \|(T - \alpha I)^{-1}\| = 0$.

(10.36) Result (Spectrum is closed). The spectrum of a bounded operator on a Banach space is a closed subset of $\mathbb{F}$.

(10.37) Result (Analyticity of $\alpha \mapsto \langle (T - \alpha I)^{-1} f, g\rangle$). Suppose V is a complex Hilbert space and $T \in B(V)$. Then for fixed $f, g \in V$, the function

$$\alpha \longmapsto \langle (T - \alpha I)^{-1} f,\ g\rangle$$

is analytic on $\mathbb{C} \setminus \operatorname{sp}(T)$.

Proof sketch. Fix $\beta \in \mathbb{C} \setminus \mathrm{sp}(T)$. If $|\alpha - \beta| < \|(T - \beta I)^{-1}\|^{-1}$, then

$$T - \alpha I = (T - \beta I)\Big(I - (\alpha - \beta)(T - \beta I)^{-1}\Big),$$

and the Neumann series gives

$$\Big(I - (\alpha - \beta)(T - \beta I)^{-1}\Big)^{-1} = \sum_{k=0}^{\infty}(\alpha - \beta)^k (T - \beta I)^{-k}.$$

Thus $(T-\alpha I)^{-1}$ admits a power-series expansion in $\alpha-\beta$, and pairing with f, g preserves analyticity.

$\square$

(10.38) Result (Spectrum is nonempty). If T is a bounded operator on a nonzero complex Hilbert space, then $\mathrm{sp}(T)$ is a nonempty subset of $\mathbb{C}$.

Continued – 10B

(10.39) Definition ($p(T)$). Suppose T is an operator on a vector space V, and p is a polynomial over $\mathbb{F}$:

$$p(z) = b_0 + b_1 z + \cdots + b_n z^n.$$

Then $p(T)$ is the operator on V defined by

$$p(T) := b_0 I + b_1 T + \cdots + b_n T^n.$$

If p, q are polynomials (with coefficients in $\mathbb{F}$), then

$$(pq)(T) = p(T)\, q(T).$$

(10.40) Result (Spectral Mapping Theorem). Suppose T is a bounded operator on a complex Banach space and p is a polynomial with complex coefficients. Then

$$\mathrm{sp}(p(T)) = p(\mathrm{sp}(T)).$$

(10.44) Definition (Self-adjoint). A bounded operator T on a Hilbert space is *self-adjoint* if $T^* = T$.

Example. Suppose (X, Σ, μ) is σ-finite and $h \in L^\infty(\mu)$. Define the multiplication operator M_h on $L^2(\mu)$ by $M_h f = hf$. Then $M_h^* = M_{\overline{h}}$, hence M_h is self-adjoint iff $h = \overline{h}$ a.e., i.e.

$$\mu\big(\{x \in X : h(x) \notin \mathbb{R}\}\big) = 0.$$

(10.46) Result ($\langle Tf, f\rangle = 0\ \forall f \Rightarrow T = 0$). Suppose V is a Hilbert space, $T \in B(V)$, and

$$\langle Tf, f\rangle = 0 \qquad \forall f \in V.$$

(i) If $\mathbb{F} = \mathbb{C}$, then $T = 0$.

(ii) If $\mathbb{F} = \mathbb{R}$ and T is self-adjoint, then $T = 0$.

(10.48) Result (Self-adjoint characterized by $\langle Tf, f\rangle$). Suppose T is a bounded operator on a complex Hilbert space V. Then T is self-adjoint iff

$$\langle Tf, f\rangle \in \mathbb{R} \qquad \forall f \in V.$$

(10.49) Result (Self-adjoint operators have real spectrum). If T is a bounded self-adjoint operator on a Hilbert space, then

$$\operatorname{sp}(T) \subseteq \mathbb{R}.$$

(10.50) Definition (Normal operator). A bounded operator T on a Hilbert space is *normal* if it commutes with its adjoint:

$$T^*T = TT^*.$$

Examples.

(i) (Normal.) If (X, Σ, μ) is a measure space, $h \in L^\infty(\mu)$, and $M_h f = hf$ on $L^2(\mu)$, then $M_h^* = M_{\overline{h}}$ and

$$M_h^* M_h = M_{|h|^2} = M_h M_h^*,$$

so M_h is normal. If h is real-valued a.e., then M_h is self-adjoint.

(ii) (Not normal.) Let T be the right shift on ℓ^2:

$$T(a_1, a_2, \dots) = (0, a_1, a_2, \dots).$$

Then T^* is the left shift, $T^*T = I$, but $TT^* \neq I$, hence $T^*T \neq TT^*$.

(10.53) Result (Normal in terms of norms). Suppose T is a bounded operator on a Hilbert space V. Then T is normal iff

$$\|Tf\| = \|T^*f\| \qquad \forall f \in V.$$

(10.54) Result (Normal iff real/imaginary parts commute). Suppose T is a bounded operator on a complex Hilbert space V.

(i) There exist unique self-adjoint operators A, B on V such that

$$T = A + iB \qquad \left(A = \frac{T + T^*}{2}, \quad B = \frac{T - T^*}{2i}\right).$$

(ii) T is normal iff $AB = BA$.

(10.55) Result (Invertibility for normal operators). Suppose V is a Hilbert space and $T \in B(V)$ is normal. Then the following are equivalent:

(a) T is invertible.

(b) T is left invertible.

(c) T is right invertible.

(d) T is surjective.

(e) T is injective and has closed range.

(f) T^*T (equivalently TT^*) is invertible.

(10.56) Result (Eigenvalues of T vs. T^* for normal T). Suppose T is normal on a Hilbert space V, $\alpha \in \mathbb{F}$, and $f \in V$. Then α is an eigenvalue of T with eigenvector f iff $\overline{\alpha}$ is an eigenvalue of T^* with the same eigenvector f:

$$Tf = \alpha f \iff T^* f = \overline{\alpha}\, f.$$

(10.57) Result (Orthogonal eigenvectors for normal operators). Eigenvectors of a normal operator corresponding to distinct eigenvalues are orthogonal.

(10.58) Definition (Isometry; unitary). Suppose T is a bounded operator on a Hilbert space V.

- T is an *isometry* if $\|Tf\| = \|f\|$ for all $f \in V$.
- T is *unitary* if $T^*T = TT^* = I$.

Examples.

- Let $T \in B(\ell^2)$ be the right shift $T(a_1, a_2, \dots) = (0, a_1, a_2, \dots)$. Then T is an isometry but not unitary (since T is not surjective, equivalently $TT^* \neq I$).
- Let $T \in B(\ell^2(\mathbb{Z}))$ be the (bilateral) shift $(Tf)(n) = f(n-1)$. Then T is an isometry and unitary.
- Suppose $(b_1, b_2, \dots)$ is a bounded sequence in $\mathbb{F}$ and define $T \in B(\ell^2)$ by

  $$T(a_1, a_2, \dots) = (b_1 a_1, b_2 a_2, \dots).$$

  Then T is an isometry iff T is unitary iff $|b_k| = 1$ for all $k \in \mathbb{Z}_+$.
- More generally, for M_h on $L^2(\mu)$ given by $M_h f = hf$, M_h is an isometry iff M_h is unitary iff

  $$\mu\big(\{x \in X : \ |h(x)| \neq 1\}\big) = 0.$$

(10.60) Result (Isometries preserve inner products). Suppose T is a bounded operator on a Hilbert space V. Then the following are equivalent:

(a) T is an isometry (i.e. $\|Tf\| = \|f\|$ for all $f \in V$).

(b) $\langle Tf, Tg \rangle = \langle f, g \rangle$ for all $f, g \in V$.

(c) $T^*T = I$.

(d) For every orthonormal family $\{e_k\}_{k\in\Gamma} \subset V$, the family $\{Te_k\}_{k\in\Gamma}$ is orthonormal.

(e) There exists some orthonormal basis $\{e_k\}_{k\in\Gamma}$ of V such that $\{Te_k\}_{k\in\Gamma}$ is orthonormal.

(10.61) Result (Unitary operators and adjoints are isometries). Suppose T is a bounded operator on a Hilbert space V. Then the following are equivalent:

(a) T is unitary.

(b) T is a surjective isometry.

(c) T and T^* are both isometries.

(d) T^* is unitary.

(e) T is invertible and $T^{-1} = T^*$.

(f) For every orthonormal basis $\{e_k\}_{k\in\Gamma}$ of V, the family $\{Te_k\}_{k\in\Gamma}$ is an orthonormal basis of V.

(g) There exists some orthonormal basis $\{e_k\}_{k\in\Gamma}$ of V such that $\{Te_k\}_{k\in\Gamma}$ is an orthonormal basis of V.

(10.62) Result (Spectrum of a unitary operator). If T is unitary on a Hilbert space, then

$$\sigma(T) \subseteq \{\alpha \in \mathbb{F} : \ |\alpha| = 1\}.$$

Proof sketch. Let $\alpha \in \mathbb{F}$ with $|\alpha| \neq 1$. Compute

$$(T - \alpha I)^*(T - \alpha I) = (T^* - \overline{\alpha} I)(T - \alpha I) = (1 + |\alpha|^2)I - (\overline{\alpha}T + \alpha T^*).$$

Factor:

$$(T - \alpha I)^*(T - \alpha I) = (1 + |\alpha|^2)\Big(I - \frac{\overline{\alpha}T + \alpha T^*}{1 + |\alpha|^2}\Big).$$

Since $\|T\| = \|T^*\| = 1$,

$$\Big\|\frac{\overline{\alpha}T + \alpha T^*}{1 + |\alpha|^2}\Big\| \leq \frac{|\alpha|\|T\| + |\alpha|\|T^*\|}{1 + |\alpha|^2} = \frac{2|\alpha|}{1 + |\alpha|^2} < 1,$$

so the bracketed operator is invertible by a Neumann series. Hence $(T - \alpha I)^*(T - \alpha I)$ is invertible, so $T - \alpha I$ is invertible, and $\alpha \notin \sigma(T)$. □

(10.65) Result (Spectrum of an isometry). Suppose T is an isometry on a Hilbert space and T is not unitary. Then

$$\sigma(T) \subseteq \{\alpha \in \mathbb{F} : \ |\alpha| \leq 1\}.$$

Exercise 3. Suppose E is a bounded subset of $\mathbb{F}$. Show that there exist a Hilbert space V and $T \in \mathcal{B}(V)$ such that the set of eigenvalues of T equals E.

Construction/solution. Let

$$V = \ell^2(E) := \Big\{x = \sum_{\alpha\in E} x_\alpha e_\alpha : \ \sum_{\alpha\in E} |x_\alpha|^2 < \infty\Big\}, \qquad \langle x, y\rangle = \sum_{\alpha\in E} x_\alpha\, \overline{y_\alpha}.$$

Then V is a Hilbert space with orthonormal basis $\{e_\alpha\}_{\alpha\in E}$. Define $T \in \mathcal{B}(V)$ by

$$Tx = \sum_{\alpha\in E} \alpha x_\alpha e_\alpha \quad (\text{i.e. } (Tx)_\alpha = \alpha x_\alpha).$$

Since E is bounded, $\exists M > 0$ such that $|\alpha| \le M$ for all $\alpha \in E$, hence

$$\|Tx\|^2 = \sum_{\alpha \in E} |\alpha x_\alpha|^2 \le M^2 \sum_{\alpha \in E} |x_\alpha|^2 = M^2 \|x\|^2,$$

so $\|T\| \le M$ and T is bounded. For $\lambda \in E$, we have $Te_\lambda = \lambda e_\lambda$, so λ is an eigenvalue. Conversely, if $Tx = \lambda x$ with $x \neq 0$, choose $\alpha_0 \in E$ with $x_{\alpha_0} \neq 0$; then $\alpha_0 x_{\alpha_0} = \lambda x_{\alpha_0}$, so $\lambda = \alpha_0 \in E$. Thus $\mathrm{Eig}(T) = E$. $\square$

Exercise 6. Suppose T is a bounded operator on a complex nonzero Banach space V.

(a) Prove that for $f \in V$ and $\varphi \in V^*$, the function

$$\alpha \longmapsto \varphi\big((T - \alpha I)^{-1} f\big)$$

is analytic on $\mathbb{C} \setminus \sigma(T)$.

Solution. Let $\rho(T) = \mathbb{C} \setminus \sigma(T)$ and fix $\alpha_0 \in \rho(T)$. Set $S := (T - \alpha_0 I)^{-1} \in \mathcal{B}(V)$. For $|\alpha - \alpha_0| < 1/\|S\|$,

$$T - \alpha I = (T - \alpha_0 I)\Big(I - (\alpha - \alpha_0)S\Big),$$

so

$$(T - \alpha I)^{-1} = \Big(I - (\alpha - \alpha_0)S\Big)^{-1} S = \sum_{n=0}^{\infty} (\alpha - \alpha_0)^n S^{n+1},$$

with norm convergence (Neumann series). Applying $\varphi(\cdot f)$ gives a power series in $(\alpha - \alpha_0)$, hence analyticity on $\rho(T)$. $\square$

(b) Prove that $\sigma(T) \neq \varnothing$.

Solution (contradiction via Liouville). Assume $\sigma(T) = \varnothing$, so $\rho(T) = \mathbb{C}$. For each $f \in V$, $\varphi \in V^*$, define the entire function

$$g_{f,\varphi}(\alpha) := \varphi\big((T - \alpha I)^{-1} f\big).$$

For $|\alpha| > \|T\|$,

$$(T - \alpha I)^{-1} = -\frac{1}{\alpha}\Big(I - \frac{T}{\alpha}\Big)^{-1} = -\frac{1}{\alpha} \sum_{n=0}^{\infty} \Big(\frac{T}{\alpha}\Big)^n,$$

so $\|(T - \alpha I)^{-1}\| \le \frac{1}{|\alpha| - \|T\|}$, hence $g_{f,\varphi}$ is bounded on $\mathbb{C}$. By Liouville, $g_{f,\varphi}$ is constant; since $g_{f,\varphi}(\alpha) \to 0$ as $|\alpha| \to \infty$, the constant is 0:

$$\varphi\big((T - \alpha I)^{-1} f\big) = 0 \quad \forall \alpha \in \mathbb{C},\ \forall f,\ \forall \varphi.$$

Fix α_0 and write $S_0 = (T - \alpha_0 I)^{-1}$. Then $\varphi(S_0 f) = 0$ for all $\varphi \in V^*$, so $S_0 f = 0$ for all f (e.g. by Hahn–Banach), hence $S_0 = 0$, impossible for an inverse. Therefore $\sigma(T) \neq \varnothing$. $\square$

Exercise 23. For a bounded operator T on a Banach space V, define

$$e^T := \sum_{k=0}^{\infty} \frac{T^k}{k!}.$$

(a) Show the series converges in $\mathcal{B}(V)$ and $\|e^T\| \le e^{\|T\|}$.

Solution. Since $\|T^k/k!\| \le \|T\|^k/k!$ and $\sum_{k=0}^{\infty} \|T\|^k/k! = e^{\|T\|}$ converges, the operator series converges absolutely (hence in norm) in $\mathcal{B}(V)$, and

$$\|e^T\| \le \sum_{k=0}^{\infty} \frac{\|T\|^k}{k!} = e^{\|T\|}.$$

□

(b) If $S, T \in \mathcal{B}(V)$ and $ST = TS$, prove $e^{S+T} = e^S e^T = e^T e^S$.

Solution. Using absolute convergence,

$$e^S e^T = \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \frac{S^k}{k!} \frac{T^m}{m!}.$$

If $ST = TS$, then for each $n = k + m$ we can group terms:

$$\sum_{k+m=n} \frac{S^k T^m}{k!\, m!} = \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} S^k T^{n-k} = \frac{(S+T)^n}{n!},$$

so $e^S e^T = \sum_{n=0}^{\infty} (S+T)^n/n! = e^{S+T}$. Symmetry gives $e^T e^S = e^{S+T}$ as well. □

(c) If T is self-adjoint on a complex Hilbert space, show e^{iT} is unitary.

Solution. Since $T^* = T$, we have $(iT)^* = -iT$, and by termwise adjoint,

$$(e^{iT})^* = \Big(\sum_{k=0}^{\infty} \frac{(iT)^k}{k!} \Big)^* = \sum_{k=0}^{\infty} \frac{((iT)^*)^k}{k!} = \sum_{k=0}^{\infty} \frac{(-iT)^k}{k!} = e^{-iT}.$$

Because iT commutes with $-iT$,

$$(e^{iT})^* e^{iT} = e^{-iT} e^{iT} = e^0 = I \quad \text{and similarly} \quad e^{iT} (e^{iT})^* = I.$$

Thus e^{iT} is unitary. □

Exercise 24. A bounded operator T on a Hilbert space is a *partial isometry* if $\|Tf\| = \|f\|$ for all $f \in (\ker T)^{\perp}$. Let $(X, \mathcal{S}, \mu)$ be σ-finite and $h \in L^{\infty}(\mu)$. Let $M_h \in \mathcal{B}(L^2(\mu))$ be the multiplication operator $M_h f = hf$. Prove:

$$M_h \text{ is a partial isometry} \quad \Longleftrightarrow \quad \exists E \in \mathcal{S} \text{ s.t. } |h| = \chi_E \;\; \mu\text{-a.e.}$$

Solution sketch. Note $M_h^* = M_{\overline{h}}$, hence $M_h^* M_h = M_{|h|^2}$. A multiplication operator is an orthogonal projection iff its symbol is $\{0,1\}$-valued a.e. Thus M_h is a partial isometry $\Longleftrightarrow$ $M_h^* M_h$ is a projection $\Longleftrightarrow$ $|h|^2 = |h|^4$ a.e. $\Longleftrightarrow$ $|h|^2 \in \{0,1\}$ a.e. $\Longleftrightarrow$ $|h| = \chi_E$ a.e. for $E = \{x : |h(x)| = 1\}$. □

(10.66) Definition (Compact operator). An operator T on a Hilbert space V is *compact* if for every bounded sequence $(f_n)_{n\geq 1} \subset V$, the sequence $(Tf_n)_{n\geq 1}$ has a convergent subsequence. The collection of compact operators on V is denoted $\mathcal{C}(V)$.

(10.67) Result (Finite-dimensional range $\Rightarrow$ compact). If $T \in \mathcal{B}(V)$ and $\operatorname{ran}(T)$ is finite-dimensional, then T is compact.

(10.68) Result (Compact operators are bounded). Every compact operator on a Hilbert space is bounded.

(10.69) Result ($\mathcal{C}(V)$ is a closed two-sided ideal of $\mathcal{B}(V)$). If V is a Hilbert space, then:

- $\mathcal{C}(V)$ is a closed subspace of $\mathcal{B}(V)$ (closed in the operator norm),
- if $T \in \mathcal{C}(V)$ and $S \in \mathcal{B}(V)$, then $ST \in \mathcal{C}(V)$ and $TS \in \mathcal{C}(V)$.

Equivalently, $\mathcal{C}(V)$ is a norm-closed two-sided ideal in $\mathcal{B}(V)$.

(10.70) Result (Compact integral operators / Hilbert–Schmidt). Suppose $(X, \mathcal{S}, \mu)$ is σ-finite and $k \in L^2(\mu \times \mu)$. Define $I_k : L^2(\mu) \to L^2(\mu)$ by

$$(I_k f)(x) = \int_X k(x,y) f(y)\, d\mu(y) \qquad (f \in L^2(\mu),\ \text{a.e. } x).$$

Then I_k is compact.

(10.73) Result (Adjoint preserves compactness). For $T \in \mathcal{B}(V)$, T is compact $\iff$ T^* is compact.

(10.74) Result (No infinite-dimensional closed subspaces in the range). If T is compact on a Hilbert space, then $\operatorname{ran}(T)$ contains no infinite-dimensional closed subspace.

(10.76) Result (Compact operators are not invertible on infinite-dimensional spaces). If V is infinite-dimensional and T is compact on V, then $0 \in \sigma(T)$.

(10.77) Result (Closed range for T^*T). If T is compact on a Hilbert space, then T^*T has closed range.

(Fredholm alternative form). If T is compact on V, $f, g \in V$, and $\alpha \in \mathbb{F} \setminus \{0\}$, then

$$(T - \alpha I)g = f \text{ has a solution } g \in V \quad \iff \quad f \perp \ker(T^* - \overline{\alpha} I).$$

(10.81) Definition (Geometric multiplicity). The *geometric multiplicity* of an eigenvalue α of T is

$$\dim \ker(T - \alpha I).$$

Equivalently, it is the dimension of the eigenspace corresponding to α.

(10.82) Result (Nonzero eigenvalues have finite multiplicity). If T is compact on a Hilbert space and $\alpha \in \mathbb{F} \setminus \{0\}$, then $\ker(T - \alpha I)$ is finite-dimensional.

(10.83) Result (Injective but not surjective). If T is injective but not surjective on a vector space, then

$$\operatorname{ran}(T) \supsetneq \operatorname{ran}(T^2) \supsetneq \operatorname{ran}(T^3) \supsetneq \cdots .$$

In particular, on a finite-dimensional space, every injective operator is surjective.

(10.85) Result (Fredholm alternative). Let T be compact on a Hilbert space and $\alpha \in \mathbb{F}\setminus\{0\}$. The following are equivalent:

(i) $\alpha \in \sigma(T)$,

(ii) α is an eigenvalue of T,

(iii) $T - \alpha I$ is not surjective.

Equivalently, exactly one of the following holds:

(1) $(T - \alpha I)f = 0$ has a nonzero solution $f \in V$;

(2) For every $g \in V$, the equation $(T - \alpha I)f = g$ has a solution $f \in V$.

(10.91) Result (Equal dimensions of null spaces). If T is compact on a Hilbert space and $\alpha \in \mathbb{F}$ with $\alpha \neq 0$, then

$$\dim \ker(T - \alpha I) = \dim \ker(T^* - \overline{\alpha} I).$$

(10.93) Result (Spectrum of a compact operator). If T is compact on a Hilbert space, then for every $\varepsilon > 0$,

$$\{\alpha \in \sigma(T) : \ |\alpha| \geq \varepsilon\} \text{ is a finite set.}$$

(In particular, the only possible accumulation point of $\sigma(T)$ is 0.)

Exercises 10C

2) **Claim.** If T is a compact operator on $L^2([0,1])$, then

$$\lim_{n\to\infty} \sqrt{n}\, \|T(x^n)\|_2 = 0,$$

where x^n denotes the function $t \mapsto t^n$ in $L^2([0,1])$.

Proof. Let $f_n(t) = t^n$. Then

$$\|f_n\|_2^2 = \int_0^1 t^{2n}\, dt = \frac{1}{2n+1} \quad \Longrightarrow \quad \|f_n\|_2 = \frac{1}{\sqrt{2n+1}}.$$

Define the normalized sequence

$$u_n := \frac{f_n}{\|f_n\|_2} = \sqrt{2n+1}\, f_n \quad \text{so that} \quad \|u_n\|_2 = 1.$$

Then

$$\sqrt{n}\, \|Tf_n\|_2 = \sqrt{n} \left\| T\left(\frac{u_n}{\sqrt{2n+1}} \right) \right\|_2 = \sqrt{\frac{n}{2n+1}}\, \|Tu_n\|_2.$$

So it suffices to show $\|Tu_n\|_2 \to 0$.

Weak convergence $u_n \rightharpoonup 0$**.** For a monomial $p(t) = t^k$ $(k \in \mathbb{Z}_{\geq 0})$,

$$\langle p, u_n \rangle = \int_0^1 t^k \sqrt{2n+1}\, t^n \, dt = \frac{\sqrt{2n+1}}{n+k+1} \xrightarrow[n\to\infty]{} 0.$$

Hence $\langle p, u_n \rangle \to 0$ for all polynomials p. Since polynomials are dense in $L^2([0,1])$ and $\{u_n\}$ is bounded, a standard density argument gives $\langle g, u_n \rangle \to 0$ for all $g \in L^2([0,1])$, i.e. $u_n \rightharpoonup 0$.

Compactness $\Rightarrow$ strong convergence of images. If $u_n \rightharpoonup 0$ and T is compact, then $\|Tu_n\|_2 \to 0$ (indeed, every subsequence of $\{Tu_n\}$ has a norm-convergent subsubsequence, and bounded linearity gives $Tu_n \rightharpoonup 0$, so 0 is the only possible norm-limit).

Therefore $\|Tu_n\|_2 \to 0$, and since $\sqrt{\frac{n}{2n+1}} \to \frac{1}{\sqrt{2}}$,

$$\sqrt{n}\, \|Tf_n\|_2 = \sqrt{\frac{n}{2n+1}}\, \|Tu_n\|_2 \longrightarrow 0.$$

□

4) **Claim.** Suppose $h \in L^\infty(\mathbb{R})$ and define $M_h \in \mathcal{B}(L^2(\mathbb{R}))$ by $M_h f = fh$. If $\|h\|_\infty > 0$, then M_h is *not* compact.

   Proof. Assume $\|h\|_\infty > 0$ and set $\varepsilon := \frac{1}{2}\|h\|_\infty > 0$. By the definition of essential supremum, the set
   $$E := \{x \in \mathbb{R} : \ |h(x)| \geq \varepsilon\}$$
   has positive Lebesgue measure. Then $L^2(E)$ is infinite-dimensional, so choose an orthonormal sequence $\{e_n\}_{n\geq 1} \subset L^2(E) \subset L^2(\mathbb{R})$. Since $\{e_n\}$ is orthonormal, $e_n \rightharpoonup 0$. But
   $$\|M_h e_n\|_2^2 = \int_{\mathbb{R}} |h(x)e_n(x)|^2\, dx \geq \int_E |h(x)|^2 |e_n(x)|^2\, dx \geq \varepsilon^2 \int_E |e_n(x)|^2\, dx = \varepsilon^2,$$
   so $\|M_h e_n\|_2 \not\to 0$. A compact operator must send weakly convergent sequences to norm convergent sequences (in particular, $e_n \rightharpoonup 0$ would imply $\|M_h e_n\| \to 0$), contradiction. Hence M_h is not compact. □

5) **Claim.** Suppose $b = (b_1, b_2, \dots) \in \ell^\infty$. Define $T : \ell^2 \to \ell^2$ by
   $$T(a_1, a_2, \dots) = (a_1 b_1, a_2 b_2, \dots).$$
   Then T is compact $\iff$ $b_n \to 0$.

   Proof. Let $\{e_n\}$ be the standard orthonormal basis of ℓ^2.

   ($\Rightarrow$) If T is compact, then $e_n \rightharpoonup 0$ implies $\|Te_n\|_2 \to 0$. But $Te_n = b_n e_n$, so $\|Te_n\|_2 = |b_n|$. Hence $b_n \to 0$.

   ($\Leftarrow$) If $b_n \to 0$, define finite-rank operators $T_N : \ell^2 \to \ell^2$ by
   $$T_N(a_1, a_2, \dots) = (a_1 b_1, \dots, a_N b_N, 0, 0, \dots).$$
   Then each T_N has finite-dimensional range and is compact. Moreover,
   $$\|(T - T_N)a\|_2^2 = \sum_{n>N} |b_n a_n|^2 \leq \left(\sup_{n>N} |b_n|\right)^2 \sum_{n>N} |a_n|^2 \leq \left(\sup_{n>N} |b_n|\right)^2 \|a\|_2^2,$$
   so $\|T - T_N\| \leq \sup_{n>N} |b_n| \to 0$. Since the compact operators are closed in operator norm, T is compact. □

6) **Claim.** Let $T \in \mathcal{B}(V)$ where V is a Hilbert space. If there is an orthonormal basis $\{e_k\}_{k\in\Gamma}$ of V such that

$$\sum_{k\in\Gamma} \|Te_k\|^2 < \infty,$$

then T is compact.

Proof. Since $\sum_{k\in\Gamma} \|Te_k\|^2 < \infty$, the set $\{k : \ Te_k \neq 0\}$ is at most countable. Relabel so $\Gamma = \mathbb{N}$. For $n \in \mathbb{N}$ define

$$T_n x := \sum_{k=1}^{n} \langle x, e_k\rangle \, Te_k.$$

Then $\mathrm{Ran}(T_n) \subseteq \mathrm{span}\{Te_1, \dots, Te_n\}$, so T_n has finite rank and is compact.

For $x \in V$,

$$(T - T_n)x = \sum_{k>n} \langle x, e_k\rangle \, Te_k,$$

hence by Cauchy–Schwarz,

$$\|(T - T_n)x\| \leq \left(\sum_{k>n} |\langle x, e_k\rangle|^2\right)^{1/2} \left(\sum_{k>n} \|Te_k\|^2\right)^{1/2} \leq \|x\| \left(\sum_{k>n} \|Te_k\|^2\right)^{1/2}.$$

Let $R_n := \left(\sum_{k>n} \|Te_k\|^2\right)^{1/2} \to 0$. Then $\|T - T_n\| \leq R_n \to 0$, so T is a norm-limit of compact operators, hence compact. $\square$

7) **Claim.** If $\{e_k\}_{k\in\Gamma}$ and $\{f_j\}_{j\in\Omega}$ are orthonormal bases of a Hilbert space V, then

$$\sum_{k\in\Gamma} \|Te_k\|^2 \; = \; \sum_{j\in\Omega} \|Tf_j\|^2.$$

(So $\sum \|Te_k\|^2$ is independent of the chosen ONB whenever it is finite.)

Proof. Assume the sums are finite (otherwise the identity is interpreted in $[0, \infty]$). Let $A := T^*T$, so A is positive and self-adjoint. Then

$$\sum_k \|Te_k\|^2 = \sum_k \langle Te_k, Te_k\rangle = \sum_k \langle e_k, Ae_k\rangle.$$

Expand each $\langle e_k, Ae_k\rangle$ in the basis $\{f_j\}$:

$$\langle e_k, Ae_k\rangle = \sum_j \langle e_k, f_j\rangle \, \langle f_j, Ae_k\rangle,$$

so (using nonnegativity to justify switching summations),

$$\sum_k \langle e_k, Ae_k\rangle = \sum_j \sum_k \langle e_k, f_j\rangle \, \langle f_j, Ae_k\rangle.$$

Since $A = A^*$, $\langle f_j, Ae_k\rangle = \langle Af_j, e_k\rangle$, and by completeness of $\{e_k\}$,

$$\sum_k \langle Af_j, e_k\rangle \, \langle e_k, f_j\rangle = \langle Af_j, f_j\rangle.$$

Thus

$$\sum_k \|Te_k\|^2 = \sum_j \langle Af_j, f_j\rangle = \sum_j \langle Tf_j, Tf_j\rangle = \sum_j \|Tf_j\|^2.$$

$\square$

Section 10D: Spectral Theorem for Compact Operators

(10.96) Result. $T^*T - \|T\|^2 I$ is not invertible. In particular, if T is bounded on a nonzero Hilbert space, then

$$\|T\|^2 \in \sigma(T^*T).$$

(10.99) Result (Self-adjoint compact operators have an eigenvalue). If T is a self-adjoint compact operator on a nonzero Hilbert space, then either $\|T\|$ or $-\|T\|$ is an eigenvalue of T.

(10.100) Definition (Invariant subspace). Let T be an operator on a vector space V. A subspace $U \subseteq V$ is *invariant* for T if

$$T(U) \subseteq U \quad \text{(equivalently: } Tf \in U \ \forall f \in U).$$

Example. For $b \in [0,1]$, the subspace

$$U_b := \{f \in L^2([0,1]) : f(t) = 0 \text{ for a.e. } t \in [b,1]\}$$

is invariant for the Volterra operator $V : L^2([0,1]) \to L^2([0,1])$ given by

$$(Vf)(x) = \int_0^x f(t)\,dt.$$

Also, for $0 \neq f \in V$, the subspace $\operatorname{span}\{f\}$ is invariant for T iff f is an eigenvector of T.

Restriction preserves compactness. If T is compact on a Hilbert space V and U is invariant for T, then $T|_U$ is compact on U.

(10.102) Result (Orthogonal complements for self-adjoint operators). If U is invariant for a self-adjoint operator T, then:

(i) $U^\perp$ is invariant for T;

(ii) $T|_{U^\perp}$ is self-adjoint on $U^\perp$.

(10.103) Result (ONB of eigenvectors ⇒ self-adjoint/normal). Suppose T is bounded on a Hilbert space V and there is an orthonormal basis of V consisting of eigenvectors of T.

(i) If V is over $\mathbb{R}$, then T is self-adjoint.

(ii) If V is over $\mathbb{C}$, then T is normal.

(10.106) Spectral theorem (self-adjoint compact operators). If T is a self-adjoint compact operator on a Hilbert space V, then:

(i) There is an orthonormal basis of V consisting of eigenvectors of T.

(ii) Equivalently, there exist a (finite or countable) index set Ω, an orthonormal family $\{e_k\}_{k\in\Omega} \subset V$, and real scalars $\{\lambda_k\}_{k\in\Omega} \subset \mathbb{R} \setminus \{0\}$ with $\lambda_k \to 0$ (if Ω is infinite) such that for all $f \in V$,

$$Tf = \sum_{k\in\Omega} \lambda_k \langle f, e_k\rangle e_k,$$

with convergence in norm (and $Tf = 0$ on $\ker(T)$).

(10.108) Spectral theorem (normal compact operators). If T is a compact operator on a complex Hilbert space V, then there is an orthonormal basis of V consisting of eigenvectors of T *iff* T is normal.

Continued 10D

(10.113) Result (Singular value decomposition). If T is a compact operator on a Hilbert space V, then there exist a (finite or countable) index set Ω, orthonormal families $\{e_k\}_{k\in\Omega}$ and $\{h_k\}_{k\in\Omega}$ in V, and positive numbers $\{s_k\}_{k\in\Omega}$ such that for all $f \in V$,

$$Tf = \sum_{k\in\Omega} s_k \langle f, e_k \rangle h_k,$$

with convergence in norm.

(10.116) Definition (Singular values). Let T be a compact operator on a Hilbert space. The *singular values* of T, denoted

$$s_1(T) \geq s_2(T) \geq s_3(T) \geq \cdots,$$

are the positive square roots of the positive eigenvalues of T^*T, listed in decreasing order with each singular value repeated according to the geometric multiplicity of the corresponding eigenvalue of T^*T. If T^*T has only finitely many positive eigenvalues, define $s_n(T) = 0$ for all remaining $n \in \mathbb{Z}^+$.

Example (finite-dimensional). Define $T : \mathbb{F}^4 \to \mathbb{F}^4$ by

$$T(z_1, z_2, z_3, z_4) = (0,\, 3z_1,\, 2z_2,\, -3z_4).$$

Then

$$(T^*T)(z_1, z_2, z_3, z_4) = (9z_1,\, 4z_2,\, 0,\, 9z_4),$$

so the eigenvalues of T^*T are $9, 4, 0$ with $\dim\ker(T^*T - 9I) = 2$ and $\dim\ker(T^*T - 4I) = 1$. Hence the singular values of T are

$$3 \geq 3 \geq 2 \geq 0 \geq 0 \geq \cdots.$$

(Note that here -3 and 0 are the only eigenvalues of T, but the singular values still detect 2 via T^*T.)

(10.120) Result (Sum of squares of singular values for integral operators). Let (X, μ) be σ-finite and suppose $k \in L^2(\mu \times \mu)$. Let T_k be the integral operator

$$(T_k f)(x) = \int_X k(x, y) f(y)\, d\mu(y).$$

Then T_k is compact and

$$\|k\|^2_{L^2(\mu\times\mu)} = \sum_{n=1}^{\infty} \big(s_n(T_k)\big)^2.$$

Proof sketch. Take an SVD $T_k f = \sum_n s_n \langle f, e_n\rangle h_n$. Extend $\{e_n\}$ to an ONB $\{e_j\}_{j\in\Gamma}$ of $L^2(\mu)$ and $\{h_n\}$ to an ONB $\{h_\ell\}_{\ell\in\Lambda}$. Define $\phi_{\ell j}(x, y) := e_j(y)\, \overline{h_\ell(x)}$. Then $\{\phi_{\ell j}\}_{\ell, j}$ is an ONB of $L^2(\mu \times \mu)$ and

$$\langle k, \phi_{\ell j}\rangle_{L^2(\mu\times\mu)} = \int_X \overline{h_\ell(x)} \left(\int_X k(x, y) e_j(y)\, d\mu(y) \right) d\mu(x) = \langle T_k e_j, h_\ell \rangle_{L^2(\mu)}.$$

By Parseval in $L^2(\mu \times \mu)$,

$$\|k\|_2^2 = \sum_{\ell,j} |\langle k, \phi_{\ell j}\rangle|^2 = \sum_{\ell,j} |\langle T_k e_j, h_\ell\rangle|^2.$$

For fixed j, Parseval in $L^2(\mu)$ gives $\sum_\ell |\langle T_k e_j, h_\ell\rangle|^2 = \|T_k e_j\|^2$, so $\|k\|_2^2 = \sum_j \|T_k e_j\|^2$. Using the SVD,

$$T_k e_j = \sum_n s_n \langle e_j, e_n\rangle h_n \quad \Rightarrow \quad \|T_k e_j\|^2 = \sum_n s_n^2 |\langle e_j, e_n\rangle|^2,$$

and summing over j plus Parseval for $\{e_j\}$ yields $\sum_j \|T_k e_j\|^2 = \sum_n s_n^2$. □

Exercises 10D

8a) **Claim.** If T is a self-adjoint compact operator on a Hilbert space, then there exists a self-adjoint compact operator S such that $S^3 = T$.

Proof. By the spectral theorem for self-adjoint compact operators,

$$Tx = \sum_n \lambda_n \langle x, e_n\rangle e_n, \qquad \lambda_n \in \mathbb{R},\ \lambda_n \to 0,$$

for some orthonormal family $\{e_n\}$. Define

$$Sx := \sum_n \sqrt[3]{\lambda_n}\, \langle x, e_n\rangle e_n.$$

Then $\sqrt[3]{\lambda_n} \in \mathbb{R}$ and $\sqrt[3]{\lambda_n} \to 0$, so S is self-adjoint and compact, and

$$S^3 x = \sum_n (\sqrt[3]{\lambda_n})^3 \langle x, e_n\rangle e_n = \sum_n \lambda_n \langle x, e_n\rangle e_n = Tx.$$

□

8b) **Claim.** If T is a normal compact operator on a complex Hilbert space, then there exists a normal compact operator S such that $S^2 = T$.

Proof. By the spectral theorem for normal compact operators,

$$Tx = \sum_n \lambda_n \langle x, e_n\rangle e_n, \qquad \lambda_n \in \mathbb{C},\ \lambda_n \to 0,$$

in some orthonormal basis of eigenvectors. Choose a branch of the square root on $\{\lambda_n\}$ and set $\mu_n = \sqrt{\lambda_n}$. Define

$$Sx := \sum_n \mu_n \langle x, e_n\rangle e_n.$$

Then S is diagonal in an ONB, hence normal; and $\mu_n \to 0$, so S is compact. Finally,

$$S^2 x = \sum_n \mu_n^2 \langle x, e_n\rangle e_n = \sum_n \lambda_n \langle x, e_n\rangle e_n = Tx.$$

□

10) **Claim.** Suppose T is a self-adjoint compact operator on a Hilbert space and $\|T\| \le \frac{1}{4}$. Then there exists a self-adjoint compact operator S such that

$$S^2 + S = T.$$

Proof. By the spectral theorem, $Tx = \sum_n \lambda_n \langle x, e_n \rangle e_n$ with $\lambda_n \in \mathbb{R}$, $\lambda_n \to 0$, and $\sigma(T) \subset [-\frac{1}{4}, \frac{1}{4}]$. Define $f : [-\frac{1}{4}, \frac{1}{4}] \to \mathbb{R}$ by

$$f(t) = \frac{-1 + \sqrt{1+4t}}{2}.$$

Then $1 + 4t \ge 0$ on this interval, f is continuous, and $f(0) = 0$. Define

$$Sx := \sum_n f(\lambda_n)\, \langle x, e_n \rangle e_n.$$

Since $f(\lambda_n) \to f(0) = 0$, S is compact, and since $f(\lambda_n) \in \mathbb{R}$, S is self-adjoint. Moreover, for each eigenvalue λ,

$$f(\lambda)^2 + f(\lambda) = \lambda,$$

so $S^2 + S = T$ on the eigenbasis, hence on all of V. $\square$

(11) For $k \in \mathbb{Z}$, define $g_k \in L^2((-\pi, \pi])$ and $h_k \in L^2((-\pi, \pi])$ by

$$g_k(t) := \frac{1}{\sqrt{2\pi}} e^{it/2} e^{ikt}, \qquad h_k(t) := \frac{1}{\sqrt{2\pi}} e^{ikt}, \qquad (\mathbb{F} = \mathbb{C}).$$

(a) *Show $\{g_k\}_{k\in\mathbb{Z}}$ is an ONB of $L^2((-\pi, \pi])$.*

For $k, j \in \mathbb{Z}$,

$$\langle g_k, g_j \rangle = \int_{-\pi}^{\pi} \frac{1}{\sqrt{2\pi}} e^{it/2} e^{ikt}\, \overline{\frac{1}{\sqrt{2\pi}} e^{it/2} e^{ijt}}\, dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(k-j)t}\, dt = \begin{cases} 1, & k = j, \\ 0, & k \ne j. \end{cases}$$

Hence $\{g_k\}$ is orthonormal. By the referenced example (density of trigonometric exponentials), $\overline{\operatorname{span}}\{g_k\} = L^2((-\pi, \pi])$, so $\{g_k\}$ is an ONB.

(b) *Use (a) to show $\{h_k\}_{k\in\mathbb{Z}}$ is an ONB of $L^2((-\pi, \pi])$.*

Define $U : L^2 \to L^2$ by

$$(Uf)(t) := e^{-it/2} f(t).$$

Since $|e^{-it/2}| = 1$ for all t, U is an isometric bijection, hence a unitary multiplication operator. Moreover,

$$(Ug_k)(t) = e^{-it/2}\Big(\frac{1}{\sqrt{2\pi}} e^{it/2} e^{ikt}\Big) = \frac{1}{\sqrt{2\pi}} e^{ikt} = h_k(t).$$

Unitary operators map ONBs to ONBs, so $\{h_k\}_{k\in\mathbb{Z}}$ is an ONB.

(c) *Use (b) to show the orthonormal family from Example 8.51 is an ONB of $L^2((-\pi, \pi])$.*

Let $\mathcal{H} := \{h_k\}_{k\in\mathbb{Z}}$ and let

$$\mathcal{T} := \Big\{\frac{1}{\sqrt{2\pi}}\Big\} \cup \Big\{\frac{\cos(kt)}{\sqrt{\pi}}, \frac{\sin(kt)}{\sqrt{\pi}} : k \ge 1\Big\}.$$

Using Euler relations and $h_k(t) = \frac{1}{\sqrt{2\pi}} e^{ikt}$,

$$\frac{\cos(kt)}{\sqrt{\pi}} = \frac{1}{\sqrt{2}}\,(h_k + h_{-k}), \qquad \frac{\sin(kt)}{\sqrt{\pi}} = \frac{1}{\sqrt{2}i}\,(h_k - h_{-k}).$$

Thus $\operatorname{span}(\mathcal{T}) = \operatorname{span}_{\mathbb{C}}(\mathcal{H})$. Since $\mathcal{H}$ is an ONB, $\overline{\operatorname{span}}(\mathcal{H}) = L^2$, hence $\overline{\operatorname{span}}(\mathcal{T}) = L^2$. Because $\mathcal{T}$ is orthonormal and its span is dense, $\mathcal{T}$ is an ONB.

(14) Suppose T is a compact operator on a Hilbert space V with singular value decomposition

$$Tf = \sum_{k=1}^{\infty} s_k(T)\,\langle f, e_k\rangle\, h_k \qquad (\forall f \in V).$$

For $n \in \mathbb{Z}_+$ define $T_n : V \to V$ by

$$T_n f := \sum_{k=1}^{n} s_k(T)\,\langle f, e_k\rangle\, h_k.$$

Show that $\lim_{n\to\infty} \|T - T_n\| = 0$.

Let $R_n := T - T_n$. Then

$$R_n f = \sum_{k=n+1}^{\infty} s_k(T)\,\langle f, e_k\rangle\, h_k.$$

Using orthonormality of $\{h_k\}$ and Parseval,

$$\|R_n f\|^2 = \sum_{k=n+1}^{\infty} |s_k(T)|^2\,|\langle f, e_k\rangle|^2 \le s_{n+1}(T)^2 \sum_{k=n+1}^{\infty} |\langle f, e_k\rangle|^2 \le s_{n+1}(T)^2 \|f\|^2,$$

where the last inequality is Bessel's inequality. Hence $\|R_n\| \le s_{n+1}(T)$. Since T is compact, $s_n(T) \to 0$, so $\|T - T_n\| = \|R_n\| \to 0$.

(16) Suppose T is compact on a Hilbert space V and $n \in \mathbb{Z}_+$. Prove

$$s_n(T) = \inf\Big\{\ \|T|_{U^\perp}\|\ :\ U \subseteq V \text{ subspace},\ \dim(U) < n\Big\}.$$

Let $\mathcal{U}_n := \{U \le V : \dim(U) < n\}$.

- *Upper bound.* Let $U_0 := \operatorname{span}\{e_1, \dots, e_{n-1}\}$, so $\dim(U_0) = n-1$ and $U_0 \in \mathcal{U}_n$. If $u \in U_0^\perp$, then $u = \sum_{j\ge n} c_j e_j$, hence

  $$Tu = \sum_{j\ge n} s_j c_j h_j, \qquad \|Tu\|^2 = \sum_{j\ge n} s_j^2 |c_j|^2 \le s_n(T)^2 \sum_{j\ge n} |c_j|^2 = s_n(T)^2 \|u\|^2.$$

  Thus $\|T|_{U_0^\perp}\| \le s_n(T)$, so $\inf_{U\in\mathcal{U}_n} \|T|_{U^\perp}\| \le s_n(T)$.

- *Lower bound.* Fix arbitrary $U \in \mathcal{U}_n$. Let $W := \operatorname{span}\{e_1, \dots, e_n\}$ so $\dim(W) = n$. Since $\dim(U) < n$, we have $\dim(W \cap U^\perp) \ge n - \dim(U) > 0$, hence choose $0 \ne v \in W \cap U^\perp$ and normalize $\|v\| = 1$. Writing $v = \sum_{j=1}^{n} c_j e_j$ gives

  $$\|Tv\|^2 = \sum_{j=1}^{n} s_j^2 |c_j|^2 \ge s_n(T)^2 \sum_{j=1}^{n} |c_j|^2 = s_n(T)^2 \|v\|^2 = s_n(T)^2.$$

  Therefore $\|T|_{U^\perp}\| \ge \|Tv\| \ge s_n(T)$. Taking the infimum over U yields $\inf_{U\in\mathcal{U}_n} \|T|_{U^\perp}\| \ge s_n(T)$.

Combining both inequalities gives the claimed equality.

(17) Suppose T is compact on a Hilbert space V with SVD

$$Tf = \sum_{k\geq 1} s_k \langle f, e_k\rangle h_k.$$

Show that

$$T^* g = \sum_{k\geq 1} s_k \langle g, h_k\rangle e_k \qquad (\forall g \in V).$$

Let $f, g \in V$. Then

$$\langle Tf, g\rangle = \Big\langle \sum_{k\geq 1} s_k \langle f, e_k\rangle h_k,\ g\Big\rangle = \sum_{k\geq 1} s_k \langle f, e_k\rangle \langle h_k, g\rangle = \sum_{k\geq 1} s_k \langle f, e_k\rangle \overline{\langle g, h_k\rangle}.$$

Using conjugate-linearity in the second argument,

$$\langle Tf, g\rangle = \Big\langle f,\ \sum_{k\geq 1} s_k \langle g, h_k\rangle e_k\Big\rangle = \langle f, T^* g\rangle.$$

Since f is arbitrary, $T^* g = \sum_{k\geq 1} s_k \langle g, h_k\rangle e_k$.

(18) Suppose T is an operator on a finite-dimensional Hilbert space V with $\dim(V) = n$.

(a) *Prove T is invertible iff $s_n(T) \neq 0$.*

Let $Tu = \sum_{j=1}^n s_j \langle u, e_j\rangle h_j$ with $s_1 \geq \cdots \geq s_n \geq 0$. If T is invertible, then $\ker(T) = \{0\}$. If $s_n(T) = 0$, then $Te_n = s_n h_n = 0$ with $e_n \neq 0$, contradicting injectivity. Hence $s_n(T) \neq 0$.

Conversely, assume $s_n(T) > 0$. For any $f \in V$,

$$\|Tf\|^2 = \sum_{j=1}^n s_j^2 |\langle f, e_j\rangle|^2 \geq s_n(T)^2 \sum_{j=1}^n |\langle f, e_j\rangle|^2 = s_n(T)^2 \|f\|^2.$$

So if $Tf = 0$ then $\|f\| = 0$, hence $f = 0$ and $\ker(T) = \{0\}$. In finite dimension, injective $\Leftrightarrow$ invertible.

(b) *Assume T is invertible and has SVD $Tf = \sum_{k=1}^n s_k \langle f, e_k\rangle h_k$. Show*

$$T^{-1} f = \sum_{k=1}^n \frac{\langle f, h_k\rangle}{s_k}\, e_k.$$

Define $S \in \mathcal{B}(V)$ by

$$Sf := \sum_{k=1}^n \frac{\langle f, h_k\rangle}{s_k}\, e_k.$$

Then for each $f \in V$,

$$STf = \sum_{k=1}^n \frac{\langle Tf, h_k\rangle}{s_k} e_k.$$

But $\langle Tf, h_k\rangle = \sum_{j=1}^n s_j \langle f, e_j\rangle \langle h_j, h_k\rangle = s_k \langle f, e_k\rangle$, so $STf = \sum_{k=1}^n \langle f, e_k\rangle e_k = f$ (since $\{e_k\}$ is an ONB). Hence $ST = I$, so $S = T^{-1}$.

(19) Suppose T is compact on a Hilbert space V. Prove that for any ONB $\{e_k\}_{k\geq 1}$ of V,

$$\sum_{k\geq 1} \|Te_k\|^2 = \sum_{n\geq 1} s_n(T)^2.$$

Let $Tf = \sum_{n\geq 1} s_n\langle f, u_n\rangle v_n$ be an SVD, with $\{u_n\}$ and $\{v_n\}$ orthonormal. For each k,

$$\|Te_k\|^2 = \sum_{n\geq 1} |\langle Te_k, v_n\rangle|^2 \quad \text{(Parseval in the ONB } \{v_n\}).$$

Moreover,

$$\langle Te_k, v_n\rangle = \sum_{m\geq 1} s_m\langle e_k, u_m\rangle\langle v_m, v_n\rangle = s_n\langle e_k, u_n\rangle.$$

Thus

$$\sum_{k\geq 1} \|Te_k\|^2 = \sum_{k\geq 1}\sum_{n\geq 1} s_n^2|\langle e_k, u_n\rangle|^2 = \sum_{n\geq 1} s_n^2 \sum_{k\geq 1} |\langle u_n, e_k\rangle|^2 = \sum_{n\geq 1} s_n^2\|u_n\|^2 = \sum_{n\geq 1} s_n^2,$$

where Tonelli/Fubini applies since all terms are nonnegative, and the inner sum equals $\|u_n\|^2 = 1$ by Parseval.

(20) Use Example 10.24 to evaluate $\sum_{n=1}^{\infty} \frac{1}{n^2}$.

Let $S := \sum_{n=1}^{\infty} \frac{1}{n^2}$ and split into odd and even parts:

$$S = \sum_{\substack{n\geq 1\\ n \text{ odd}}} \frac{1}{n^2} + \sum_{\substack{n\geq 1\\ n \text{ even}}} \frac{1}{n^2}.$$

The even part is

$$\sum_{n \text{ even}} \frac{1}{n^2} = \sum_{k=1}^{\infty} \frac{1}{(2k)^2} = \frac{1}{4}\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{1}{4}S.$$

Hence $\sum_{n \text{ odd}} \frac{1}{n^2} = S - \frac{1}{4}S = \frac{3}{4}S$. From Example 10.24, $\sum_{n \text{ odd}} \frac{1}{n^2} = \frac{\pi^2}{8}$, so

$$\frac{3}{4}S = \frac{\pi^2}{8} \quad \Longrightarrow \quad S = \frac{\pi^2}{6}.$$

11A. Fourier Series and Poisson Integral

Fourier coefficients and the Riemann–Lebesgue lemma.

- For $k \in \mathbb{Z}$, define $e_k : (-\pi, \pi] \to \mathbb{R}$ by

$$e_k(t) = \begin{cases} \frac{1}{\sqrt{\pi}}\sin(kt), & k > 0,\\ \frac{1}{\sqrt{2\pi}}, & k = 0,\\ \frac{1}{\sqrt{\pi}}\cos(kt), & k < 0. \end{cases}$$

  Then $\{e_k\}_{k\in\mathbb{Z}}$ is an orthonormal basis of $L^2((-\pi, \pi])$.

- In this chapter we work on the unit circle in $\mathbb{C}$ instead of the interval $(-\pi, \pi]$ via the map

$$t \longmapsto e^{it} = \cos t + i\sin t.$$

(11.3) Definition (Unit disk and unit circle).

$$\mathbb{D} := \{w \in \mathbb{C} : \ |w| < 1\}, \qquad \partial\mathbb{D} := \{z \in \mathbb{C} : \ |z| = 1\}.$$

(11.4) Definition (Measurable subsets of $\partial\mathbb{D}$ and the measure σ).

- A set $E \subseteq \partial\mathbb{D}$ is *measurable* if $\{t \in (-\pi, \pi] : \ e^{it} \in E\}$ is a Borel subset of $\mathbb{R}$.
- Define σ on measurable subsets of $\partial\mathbb{D}$ by transferring (normalised) Lebesgue measure from $(-\pi, \pi]$:
$$\sigma(E) := \frac{1}{2\pi}\, m\big(\{t \in (-\pi, \pi] : \ e^{it} \in E\}\big), \qquad \text{so that } \sigma(\partial\mathbb{D}) = 1.$$
- This definition gives the change-of-variables formula
$$\int_{\partial\mathbb{D}} f(z)\, d\sigma(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{it})\, dt, \qquad f : \partial\mathbb{D} \to \mathbb{C} \text{ measurable.}$$

(11.5) Definition ($L^p(\partial\mathbb{D})$). For $1 \le p \le \infty$, define $L^p(\partial\mathbb{D})$ as the usual complex L^p-space over $(\partial\mathbb{D}, \sigma)$. Note: if $z = e^{it}$ then $\overline{z} = e^{-it}$ and $z^n \overline{z}^m = z^{n-m}$.

(11.6) Result (Orthonormal family in $L^2(\partial\mathbb{D})$). The family $\{z^n\}_{n\in\mathbb{Z}}$ is orthonormal in $L^2(\partial\mathbb{D})$:

$$\langle z^m, z^n \rangle = \int_{\partial\mathbb{D}} z^m\, \overline{z^n}\, d\sigma = \int_{\partial\mathbb{D}} z^{m-n}\, d\sigma = \begin{cases} 1, & m = n, \\ 0, & m \ne n. \end{cases}$$

(Equivalently, $\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(m-n)t}\, dt$.)

Moreover, if $f \in \overline{\operatorname{span}}\{z^n : n \in \mathbb{Z}\} \subseteq L^2(\partial\mathbb{D})$, then

$$f = \sum_{n\in\mathbb{Z}} \langle f, z^n \rangle z^n$$

with convergence in $L^2(\partial\mathbb{D})$ (unconditional in the Hilbert space sense).

(11.7) Definition (Fourier coefficient and Fourier series). For $f \in L^1(\partial\mathbb{D})$ (in particular for $f \in L^2(\partial\mathbb{D})$) and $n \in \mathbb{Z}$, define

$$\widehat{f}(n) := \int_{\partial\mathbb{D}} f(z)\, \overline{z^n}\, d\sigma(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{it}) e^{-int}\, dt.$$

The *Fourier series* of f is the formal sum

$$\sum_{n=-\infty}^{\infty} \widehat{f}(n)\, z^n.$$

(11.8) Examples (Fourier coefficients).

- If h is analytic on an open set containing $\overline{\mathbb{D}}$, then

$$h(z) = \sum_{n=0}^{\infty} a_n z^n \quad \text{uniformly on } \overline{\mathbb{D}}.$$

  In particular $h|_{\partial\mathbb{D}} \in L^2(\partial\mathbb{D})$ and its Fourier coefficients satisfy

$$\widehat{h}(n) = \begin{cases} a_n, & n \geq 0, \\ 0, & n < 0. \end{cases}$$

  (So on $\partial\mathbb{D}$, the Fourier series agrees with the Taylor series.)

- Let $f : \partial\mathbb{D} \to \mathbb{R}$ be given by

$$f(z) = \frac{1}{(3-z)(3-\overline{z})} = \frac{1}{|3-z|^2}.$$

  For $z \in \partial\mathbb{D}$ one can rewrite

$$f(z) = \frac{1}{8}\Big(\frac{z}{3-z} + \frac{\overline{z}}{3-\overline{z}}\Big) = \frac{1}{8}\Big(\frac{z/3}{1-z/3} + \frac{\overline{z}/3}{1-\overline{z}/3}\Big),$$

  and expand each term as a geometric series to obtain

$$f(z) = \frac{1}{8}\sum_{n\in\mathbb{Z}} \frac{z^n}{3^{|n|}} \qquad (z \in \partial\mathbb{D}).$$

  Hence

$$\widehat{f}(n) = \frac{1}{8} \cdot \frac{1}{3^{|n|}}, \qquad \forall n \in \mathbb{Z}.$$

(11.9) Result (Algebraic properties of Fourier coefficients). If $f, g \in L^1(\partial\mathbb{D})$ and $n \in \mathbb{Z}$, then

(i) $\widehat{(f+g)}(n) = \widehat{f}(n) + \widehat{g}(n)$,

(ii) $\widehat{(\alpha f)}(n) = \alpha\,\widehat{f}(n)$ for all $\alpha \in \mathbb{C}$,

(iii) $|\widehat{f}(n)| \leq \|f\|_1$.

(11.10) Result (Riemann–Lebesgue lemma). If $f \in L^1(\partial\mathbb{D})$, then $\widehat{f}(n) \to 0$ as $|n| \to \infty$.

Proof sketch. Fix $\varepsilon > 0$. Choose $g \in L^2(\partial\mathbb{D}) \cap L^1(\partial\mathbb{D})$ such that $\|f - g\|_1 < \varepsilon$. By Bessel's inequality, $\sum_{n\in\mathbb{Z}} |\widehat{g}(n)|^2 \leq \|g\|_2^2 < \infty$, so $|\widehat{g}(n)| < \varepsilon$ for all $|n| \geq N$. Then for $|n| \geq N$,

$$|\widehat{f}(n)| \leq |\widehat{f}(n) - \widehat{g}(n)| + |\widehat{g}(n)| \leq \|f - g\|_1 + \varepsilon < 2\varepsilon,$$

so $\widehat{f}(n) \to 0$.

(11.11) Definition (Poisson-type operator via Fourier series). For $f \in L^1(\partial\mathbb{D})$ and $0 \leq r < 1$, define $P_r f : \partial\mathbb{D} \to \mathbb{C}$ by

$$(P_r f)(z) := \sum_{n=-\infty}^{\infty} r^{|n|}\, \widehat{f}(n)\, z^n.$$

There are no convergence issues since $|z^n| = 1$ on $\partial\mathbb{D}$ and

$$\sum_{n\in\mathbb{Z}} \left|r^{|n|}\widehat{f}(n)z^n\right| \leq \|f\|_1 \sum_{n\in\mathbb{Z}} r^{|n|} = \|f\|_1\Big(1 + 2\sum_{n\geq 1} r^n\Big) = \|f\|_1 \frac{1+r}{1-r} < \infty.$$

Hence the partial sums converge uniformly on $\partial\mathbb{D}$, and $P_r f$ is continuous on $\partial\mathbb{D}$.

(II.14) Definition (Poisson kernel). For $0 < r < 1$, define $P_r : \partial\mathbb{D} \to (0,\infty)$ by

$$P_r(\zeta) \;:=\; \frac{1-r^2}{|1-r\zeta|^2}, \qquad \zeta \in \partial\mathbb{D}.$$

The family $\{P_r\}_{0<r<1}$ is called the *Poisson kernel* on $\mathbb{D}$.

(II.15) Result (Integral formula / Poisson averaging). Let $f \in L^1(\partial\mathbb{D})$ and $0 < r < 1$. For $z \in \partial\mathbb{D}$,

$$(P_r f)(z) \;=\; \int_{\partial\mathbb{D}} f(\omega)\, P_r(z\overline{\omega})\, d\sigma(\omega) \;=\; \int_{\partial\mathbb{D}} f(\omega)\, \frac{1-r^2}{|1-rz\overline{\omega}|^2}\, d\sigma(\omega),$$

where $d\sigma$ is normalized arc-length measure on $\partial\mathbb{D}$ (equivalently $d\sigma(\mathrm{e}^{it}) = \frac{dt}{2\pi}$).

(II.16) Result (Approximate identity properties of P_r). The Poisson kernels satisfy:

(a) $P_r(\zeta) > 0$ for all $r \in (0,1)$ and $\zeta \in \partial\mathbb{D}$.

(b) $\displaystyle\int_{\partial\mathbb{D}} P_r(\zeta)\, d\sigma(\zeta) = 1$ for each $r \in (0,1)$.

(c) For every $\delta > 0$,

$$\lim_{r\uparrow 1} \int_{\{\,\mathrm{e}^{it}:\ |t|\geq\delta\,\}} P_r(\mathrm{e}^{it})\, \frac{dt}{2\pi} = 0,$$

i.e. the mass concentrates near 1 as $r \uparrow 1$.

(II.18) Result (Uniform approximation for continuous data). If $f : \partial\mathbb{D} \to \mathbb{C}$ is continuous, then

$$\lim_{r\uparrow 1} \|f - P_r f\|_\infty = 0.$$

Equivalently, $P_r f \to f$ uniformly on $\partial\mathbb{D}$ as $r \uparrow 1$.

(II.19) Definition (Harmonic function). Let $G \subset \mathbb{R}^2$ be open. A function $u : G \to \mathbb{C}$ is *harmonic* if

$$\frac{\partial^2 u}{\partial x^2}(w) + \frac{\partial^2 u}{\partial y^2}(w) = 0, \qquad \forall w \in G.$$

The left-hand side is the *Laplacian* $\Delta u(w)$; thus u is harmonic iff $\Delta u \equiv 0$ on G.

(II.22) Result (Poisson integral is harmonic). Let $f \in L^1(\partial\mathbb{D})$. Define $u : \mathbb{D} \to \mathbb{C}$ by

$$u(rz) \;:=\; (P_r f)(z), \qquad 0 < r < 1,\; z \in \partial\mathbb{D}.$$

Then u is harmonic on $\mathbb{D}$. This u is called the *Poisson integral* of f. Moreover, if $w = rz \in \mathbb{D}$ ($z \in \partial\mathbb{D}$), then (in terms of Fourier coefficients)

$$u(w) = \sum_{n\in\mathbb{Z}} \widehat{f}(n)\, r^{|n|} z^n = \sum_{n\geq 0} \widehat{f}(n)\, w^n \;+\; \sum_{n\geq 1} \widehat{f}(-n)\, \overline{w}^{\,n}.$$

(II.23) Result (Dirichlet problem on $\mathbb{D}$). If $f : \partial\mathbb{D} \to \mathbb{C}$ is continuous, define $u : \overline{\mathbb{D}} \to \mathbb{C}$ by

$$u(rz) = (P_r f)(z) \quad (0 \leq r < 1,\; z \in \partial\mathbb{D}), \qquad \text{and} \qquad u(z) = f(z) \quad (z \in \partial\mathbb{D}).$$

Then u is continuous on $\overline{\mathbb{D}}$, harmonic on $\mathbb{D}$, and $u|_{\partial\mathbb{D}} = f$.

(II.24) Definition (k times continuously differentiable on $\partial\mathbb{D}$). Let $k \in \mathbb{Z}_{\geq 0}$ and $f : \partial\mathbb{D} \to \mathbb{C}$. Define the 2π-periodic lift

$$\widetilde{f} : \mathbb{R} \to \mathbb{C}, \qquad \widetilde{f}(t) := f(\mathrm{e}^{it}).$$

We say f is *k times continuously differentiable* (write $f \in C^k(\partial\mathbb{D})$) if $\widetilde{f}$ is k times differentiable on $\mathbb{R}$ and $\widetilde{f}^{(k)}$ is continuous. If $f \in C^k(\partial\mathbb{D})$, define $f^{(k)} : \partial\mathbb{D} \to \mathbb{C}$ by

$$f^{(k)}(\mathrm{e}^{it}) \;:=\; \widetilde{f}^{(k)}(t), \qquad \text{with } f^{(0)} = f.$$

(Heuristic: lift to $\mathbb{R}$, differentiate, then push back to $\partial\mathbb{D}$.)

(II.26) Result (Fourier coefficients of derivatives). If $k \in \mathbb{Z}_{>0}$ and $f \in C^k(\partial\mathbb{D})$, then for every $n \in \mathbb{Z}$,

$$\widehat{f^{(k)}}(n) = (in)^k\, \widehat{f}(n).$$

(Proof sketch: integration by parts on $\mathbb{R}$ using periodicity.)

(II.27) Result (Uniform convergence for C^2 data). If $f \in C^2(\partial\mathbb{D})$, then its Fourier series converges uniformly and equals f:

$$f(z) = \sum_{n\in\mathbb{Z}} \widehat{f}(n)\, z^n, \qquad z \in \partial\mathbb{D},$$

and the symmetric partial sums $\sum_{n=-N}^{N} \widehat{f}(n) z^n$ converge uniformly on $\partial\mathbb{D}$. (Reason: from (II.26) with $k = 2$, $|\widehat{f}(n)| \lesssim 1/n^2$, so the series is absolutely and uniformly convergent.)

Exercise (1). For $f \in L^1(\partial\mathbb{D})$ and $n \in \mathbb{Z}$,

$$\widehat{\overline{f}}(n) = \overline{\widehat{f}(-n)}.$$

(Compute directly from the definition and take complex conjugates.)

Exercise (2). Fix $1 \le p \le \infty$ and $n \in \mathbb{Z}$. Define $\Lambda_n : L^p(\partial\mathbb{D}) \to \mathbb{C}$ by $\Lambda_n(f) = \widehat{f}(n)$.

(a) Λ_n is bounded linear and $\|\Lambda_n\| = 1$.

$$|\widehat{f}(n)| = \left| \int_{\partial\mathbb{D}} f(\zeta)\, \overline{\zeta}^{\,n}\, d\sigma(\zeta) \right| \le \|f\|_p \, \|\overline{\zeta}^{\,n}\|_q = \|f\|_p,$$

where q is the Hölder conjugate of p (and $\|\overline{\zeta}^{\,n}\|_q = 1$). Equality is attained e.g. by $f(\zeta) = \zeta^n$ (normalized), so $\|\Lambda_n\| = 1$.

(b) Characterize the extremizers: find $f \in L^p(\partial\mathbb{D})$ with $\|f\|_p = 1$ and $|\widehat{f}(n)| = 1$. Equality in Hölder forces $|f|$ to be a.e. constant and the argument aligned with ζ^n; thus (up to a unimodular constant)

$$f(\zeta) = c\,\zeta^n \quad \text{a.e. on } \partial\mathbb{D}, \qquad |c| = 1.$$

Exercise (3). For $0 \le r < 1$ and $t \in \mathbb{R}$,

$$P_r(\mathrm{e}^{it}) = \frac{1 - r^2}{1 - 2r\cos t + r^2}.$$

(Use $|1 - re^{it}|^2 = (1 - re^{it})(1 - re^{-it}) = 1 - 2r\cos t + r^2$.)

Exercise (4). If $f \in L^1(\partial\mathbb{D})$, $z \in \partial\mathbb{D}$, and f is continuous at z, then

$$\lim_{r\uparrow 1} (P_r f)(z) = f(z).$$

(Use the approximate identity properties: split the integral into a small arc near z and its complement.)

Exercise (5). Assume $f \in L^1(\partial\mathbb{D})$, $z \in \partial\mathbb{D}$, and the one-sided limits exist:

$$\lim_{t\downarrow 0} f(\mathrm{e}^{it} z) = a, \qquad \lim_{t\uparrow 0} f(\mathrm{e}^{it} z) = b.$$

Then

$$\lim_{r\uparrow 1} (P_r f)(z) = \frac{a+b}{2}.$$

(Proof idea: split the integral into $t \in (0,\delta)$, $t \in (-\delta, 0)$, and the outer region; use that $P_r(\mathrm{e}^{it})$ is even in t and concentrates at $t = 0$.)

Exercise (6). For every $p \in [1, \infty)$, there exists $f \in L^1(\partial\mathbb{D})$ such that

$$\sum_{n=1}^{\infty} |\widehat{f}(n)|^p = \infty.$$

Construction: let $\alpha = \frac{1}{p} \in (0, 1]$ and prescribe Fourier coefficients

$$\widehat{f}(n) = n^{-\alpha} \ (n \ge 1), \qquad \widehat{f}(n) = 0 \ (n \le 0).$$

Then $\sum_{n\ge 1} |\widehat{f}(n)|^p = \sum_{n \ge 1} n^{-1} = \infty$. Using standard asymptotics for trigonometric series with monotone coefficients (here $n^{-\alpha}$), one gets $f(\mathrm{e}^{it}) = O(|t|^{\alpha - 1})$ as $t \to 0$, which is integrable for $\alpha > 0$, hence $f \in L^1(\partial\mathbb{D})$.

Exercise (9) (worked from Example 11.8-style computation). Let

$$f(\mathrm{e}^{it}) = \frac{8}{|3 - 2\mathrm{e}^{it}|^2} = \frac{8}{13 - 12\cos t} \qquad (t \in \mathbb{R}).$$

Using the known Fourier expansion of $\frac{1}{13-12\cos t}$ (with ratio $2/3$), one obtains

$$f(\mathrm{e}^{it}) = \frac{8}{5}\sum_{n\in\mathbb{Z}} \left(\frac{2}{3}\right)^{|n|} \mathrm{e}^{int}.$$

Hence the Poisson integral scales coefficients by $r^{|n|}$:

$$(Pf)(re^{it}) = \frac{8}{5}\sum_{n\in\mathbb{Z}} \left(\frac{2r}{3}\right)^{|n|} \mathrm{e}^{int} = \frac{8}{5} P_{2r/3}(\mathrm{e}^{it}) = \frac{8}{5} \cdot \frac{9 - 4r^2}{|3 - 2r\mathrm{e}^{it}|^2}.$$

In particular, $(Pf)(\mathrm{e}^{it}) = f(\mathrm{e}^{it})$.

Exercise (10). If $f \in C^3(\partial\mathbb{D})$, then the Fourier series may be differentiated term-by-term:

$$f'(z) = \sum_{n\in\mathbb{Z}} (in)\, \widehat{f}(n)\, z^n, \qquad z \in \partial\mathbb{D},$$

with uniform convergence on $\partial\mathbb{D}$. (Reason: from (II.26) with $k = 3$, $|n|^3|\widehat{f}(n)| \lesssim 1$, so $\sum |n||\widehat{f}(n)| < \infty$ and the derivative series converges uniformly.)

Exercise (11) (Dirichlet kernel / unbounded partial sum functionals). Let $C(\partial\mathbb{D})$ be the Banach space of continuous functions $\partial\mathbb{D} \to \mathbb{C}$ with $\|\cdot\|_\infty$. For $M \in \mathbb{Z}_{\geq 1}$ define $\varphi_M : C(\partial\mathbb{D}) \to \mathbb{C}$ by

$$\varphi_M(f) := \sum_{n=-M}^{M} \widehat{f}(n),$$

i.e. the symmetric partial sum evaluated at $z = 1$.

(a) (Dirichlet kernel formula.) Let

$$D_M(t) := \sum_{n=-M}^{M} \mathrm{e}^{int} = \frac{\sin\big((M + \frac{1}{2})t\big)}{\sin(t/2)}.$$

Then

$$\varphi_M(f) = \int_{-\pi}^{\pi} f(\mathrm{e}^{it})\, D_M(t)\, \frac{dt}{2\pi}, \qquad \forall f \in C(\partial\mathbb{D}).$$

(b) $\displaystyle\lim_{M\to\infty} \int_{-\pi}^{\pi} |D_M(t)|\, \frac{dt}{2\pi} = \infty$. (Indeed the L^1-norm of D_M grows like $\log M$; a standard lower bound comes from comparing $\sin(t/2)$ with $t/2$ on small intervals and summing a harmonic series.)

(c) $\displaystyle\sup_{M\geq 1} \|\varphi_M\| = \infty$. (Using (a), $\|\varphi_M\| \geq \int |D_M|\, dt/(2\pi)$ by approximating $\operatorname{sgn}(D_M)$ with continuous functions.)

(d) There exists $f \in C(\partial\mathbb{D})$ such that $\{\varphi_M(f)\}_{M\geq 1}$ diverges (equivalently, the Fourier series of f diverges at $z = 1$). (If every f had convergent partial sums at 1, then $\{\varphi_M\}$ would be pointwise bounded; the Uniform Boundedness Principle would force $\sup_M \|\varphi_M\| < \infty$, contradicting (c).)

(II.14) Definition (Poisson kernel). For $0 < r < 1$, define $P_r : \partial\mathbb{D} \to (0, \infty)$ by

$$P_r(\zeta) \;:=\; \frac{1-r^2}{|1-r\zeta|^2}, \qquad \zeta \in \partial\mathbb{D}.$$

The family $\{P_r\}_{0<r<1}$ is called the *Poisson kernel* on $\mathbb{D}$.

(II.15) Result (Integral formula / Poisson averaging). Let $f \in L^1(\partial\mathbb{D})$ and $0 < r < 1$. For $z \in \partial\mathbb{D}$,

$$(P_r f)(z) \;=\; \int_{\partial\mathbb{D}} f(\omega)\, P_r(z\overline{\omega})\, d\sigma(\omega) \;=\; \int_{\partial\mathbb{D}} f(\omega)\, \frac{1-r^2}{|1-rz\overline{\omega}|^2}\, d\sigma(\omega),$$

where $d\sigma$ is normalized arc-length measure on $\partial\mathbb{D}$ (equivalently $d\sigma(\mathrm{e}^{it}) = \frac{dt}{2\pi}$).

(II.16) Result (Approximate identity properties of P_r). The Poisson kernels satisfy:

(a) $P_r(\zeta) > 0$ for all $r \in (0,1)$ and $\zeta \in \partial\mathbb{D}$.

(b) $\displaystyle\int_{\partial\mathbb{D}} P_r(\zeta)\, d\sigma(\zeta) = 1$ for each $r \in (0,1)$.

(c) For every $\delta > 0$,

$$\lim_{r\uparrow 1} \int_{\{\,\mathrm{e}^{it}:\ |t|\geq\delta\,\}} P_r(\mathrm{e}^{it})\, \frac{dt}{2\pi} = 0,$$

i.e. the mass concentrates near 1 as $r \uparrow 1$.

(II.18) Result (Uniform approximation for continuous data). If $f : \partial\mathbb{D} \to \mathbb{C}$ is continuous, then

$$\lim_{r\uparrow 1} \|f - P_r f\|_\infty = 0.$$

Equivalently, $P_r f \to f$ uniformly on $\partial\mathbb{D}$ as $r \uparrow 1$.

(II.19) Definition (Harmonic function). Let $G \subset \mathbb{R}^2$ be open. A function $u : G \to \mathbb{C}$ is *harmonic* if

$$\frac{\partial^2 u}{\partial x^2}(w) + \frac{\partial^2 u}{\partial y^2}(w) = 0, \qquad \forall w \in G.$$

The left-hand side is the *Laplacian* $\Delta u(w)$; thus u is harmonic iff $\Delta u \equiv 0$ on G.

(II.22) Result (Poisson integral is harmonic). Let $f \in L^1(\partial\mathbb{D})$. Define $u : \mathbb{D} \to \mathbb{C}$ by

$$u(rz) \;:=\; (P_r f)(z), \qquad 0 < r < 1,\ z \in \partial\mathbb{D}.$$

Then u is harmonic on $\mathbb{D}$. This u is called the *Poisson integral* of f. Moreover, if $w = rz \in \mathbb{D}$ ($z \in \partial\mathbb{D}$), then (in terms of Fourier coefficients)

$$u(w) = \sum_{n\in\mathbb{Z}} \widehat{f}(n)\, r^{|n|} z^n = \sum_{n\geq 0} \widehat{f}(n)\, w^n \;+\; \sum_{n\geq 1} \widehat{f}(-n)\, \overline{w}^{\,n}.$$

(II.23) Result (Dirichlet problem on $\mathbb{D}$). If $f : \partial\mathbb{D} \to \mathbb{C}$ is continuous, define $u : \overline{\mathbb{D}} \to \mathbb{C}$ by

$$u(rz) = (P_r f)(z) \quad (0 \le r < 1,\ z \in \partial\mathbb{D}), \qquad \text{and} \qquad u(z) = f(z) \quad (z \in \partial\mathbb{D}).$$

Then u is continuous on $\overline{\mathbb{D}}$, harmonic on $\mathbb{D}$, and $u|_{\partial\mathbb{D}} = f$.

(II.24) Definition (k times continuously differentiable on $\partial\mathbb{D}$). Let $k \in \mathbb{Z}_{\ge 0}$ and $f : \partial\mathbb{D} \to \mathbb{C}$. Define the 2π-periodic lift

$$\widetilde{f} : \mathbb{R} \to \mathbb{C}, \qquad \widetilde{f}(t) := f(\mathrm{e}^{it}).$$

We say f is *k times continuously differentiable* (write $f \in C^k(\partial\mathbb{D})$) if $\widetilde{f}$ is k times differentiable on $\mathbb{R}$ and $\widetilde{f}^{(k)}$ is continuous. If $f \in C^k(\partial\mathbb{D})$, define $f^{(k)} : \partial\mathbb{D} \to \mathbb{C}$ by

$$f^{(k)}(\mathrm{e}^{it}) \ := \ \widetilde{f}^{(k)}(t), \qquad \text{with } f^{(0)} = f.$$

(Heuristic: lift to $\mathbb{R}$, differentiate, then push back to $\partial\mathbb{D}$.)

(II.26) Result (Fourier coefficients of derivatives). If $k \in \mathbb{Z}_{>0}$ and $f \in C^k(\partial\mathbb{D})$, then for every $n \in \mathbb{Z}$,

$$\widehat{f^{(k)}}(n) = (in)^k \, \widehat{f}(n).$$

(Proof sketch: integration by parts on $\mathbb{R}$ using periodicity.)

(II.27) Result (Uniform convergence for C^2 data). If $f \in C^2(\partial\mathbb{D})$, then its Fourier series converges uniformly and equals f:

$$f(z) = \sum_{n\in\mathbb{Z}} \widehat{f}(n)\, z^n, \qquad z \in \partial\mathbb{D},$$

and the symmetric partial sums $\sum_{n=-N}^{N} \widehat{f}(n) z^n$ converge uniformly on $\partial\mathbb{D}$. (Reason: from (II.26) with $k = 2$, $|\widehat{f}(n)| \lesssim 1/n^2$, so the series is absolutely and uniformly convergent.)

Exercise (1). For $f \in L^1(\partial\mathbb{D})$ and $n \in \mathbb{Z}$,

$$\widehat{\overline{f}}(n) = \overline{\widehat{f}(-n)}.$$

(Compute directly from the definition and take complex conjugates.)

Exercise (2). Fix $1 \le p \le \infty$ and $n \in \mathbb{Z}$. Define $\Lambda_n : L^p(\partial\mathbb{D}) \to \mathbb{C}$ by $\Lambda_n(f) = \widehat{f}(n)$.

(a) Λ_n is bounded linear and $\|\Lambda_n\| = 1$.

$$|\widehat{f}(n)| = \left| \int_{\partial\mathbb{D}} f(\zeta)\, \overline{\zeta}^{\,n}\, d\sigma(\zeta) \right| \le \|f\|_p \, \|\overline{\zeta}^{\,n}\|_q = \|f\|_p,$$

where q is the Hölder conjugate of p (and $\|\overline{\zeta}^{\,n}\|_q = 1$). Equality is attained e.g. by $f(\zeta) = \zeta^n$ (normalized), so $\|\Lambda_n\| = 1$.

(b) Characterize the extremizers: find $f \in L^p(\partial\mathbb{D})$ with $\|f\|_p = 1$ and $|\widehat{f}(n)| = 1$. Equality in Hölder forces $|f|$ to be a.e. constant and the argument aligned with ζ^n; thus (up to a unimodular constant)

$$f(\zeta) = c\,\zeta^n \quad \text{a.e. on } \partial\mathbb{D}, \qquad |c| = 1.$$

Exercise (3). For $0 \le r < 1$ and $t \in \mathbb{R}$,

$$P_r(\mathrm{e}^{it}) = \frac{1-r^2}{1-2r\cos t + r^2}.$$

(Use $|1-re^{it}|^2 = (1-re^{it})(1-re^{-it}) = 1-2r\cos t + r^2$.)

Exercise (4). If $f \in L^1(\partial\mathbb{D})$, $z \in \partial\mathbb{D}$, and f is continuous at z, then

$$\lim_{r\uparrow 1}(P_r f)(z) = f(z).$$

(Use the

(2) A boundary function with jump discontinuities

Define $f : \partial\mathbb{D} \to \mathbb{R}$ by

$$f(z) = \begin{cases} 1, & \Im z > 0, \\ 0, & \Im z = 0, \\ -1, & \Im z < 0. \end{cases}$$

(Equivalently, writing $z = e^{it}$ with $t \in [-\pi, \pi]$, we have $f(e^{it}) = 1$ for $t \in (0, \pi)$ and $f(e^{it}) = -1$ for $t \in (-\pi, 0)$.)

(a) **Fourier coefficients.** For $n \in \mathbb{Z}$,

$$\widehat{f}(n) = \begin{cases} -\dfrac{2i}{\pi n}, & n \text{ odd}, \\ 0, & n \text{ even}. \end{cases}$$

Sketch of computation: for $n \neq 0$,

$$\widehat{f}(n) = \frac{1}{2\pi}\left(\int_0^\pi e^{-int}\,dt - \int_{-\pi}^0 e^{-int}\,dt\right) = \frac{1}{2\pi} \cdot \frac{2-2\cos(n\pi)}{in},$$

which vanishes for even n and equals $-2i/(\pi n)$ for odd n; also $\widehat{f}(0) = 0$.

(b) **Closed form for the Poisson integral.** For $0 < r < 1$ and $z \in \partial\mathbb{D}$,

$$(P_r f)(z) = \frac{2}{\pi}\arctan\left(\frac{2r\,\Im z}{1-r^2}\right).$$

Derivation idea: using the Fourier series form

$$(P_r f)(e^{it}) = \sum_{n\in\mathbb{Z}} \widehat{f}(n)\, r^{|n|} e^{int},$$

and combining $\pm n$ for odd n gives

$$(P_r f)(e^{it}) = \frac{4}{\pi}\sum_{k=0}^{\infty} \frac{r^{2k+1}}{2k+1}\sin\big((2k+1)t\big).$$

Recognize $\sum_{k\geq 0} \frac{w^{2k+1}}{2k+1} = \tanh^{-1}(w) = \frac{1}{2}\log\frac{1+w}{1-w}$ with $w = re^{it}$, so the sum is an argument (imaginary part of a log), yielding

$$(P_r f)(e^{it}) = \frac{2}{\pi}\operatorname{Arg}\left(\frac{1+re^{it}}{1-re^{it}}\right) = \frac{2}{\pi}\arctan\left(\frac{2r\sin t}{1-r^2}\right) = \frac{2}{\pi}\arctan\left(\frac{2r\,\Im(e^{it})}{1-r^2}\right).$$

(c) **Pointwise boundary convergence.** Fix $z \in \partial\mathbb{D}$ and write $y = \Im z$. Then as $r \to 1^-$,

$$\frac{2r\,y}{1-r^2} \to \begin{cases} +\infty, & y > 0, \\ -\infty, & y < 0, \\ 0, & y = 0, \end{cases}$$

so $\arctan(\cdot) \to \pm\frac{\pi}{2}$ (or 0), hence $(P_r f)(z) \to f(z)$ for every $z \in \partial\mathbb{D}$.

(d) **Not uniform convergence on $\partial\mathbb{D}$.** Each $P_r f$ is continuous on $\partial\mathbb{D}$ (for fixed $r < 1$), but f is discontinuous at $z = \pm 1$. A uniform limit of continuous functions is continuous, so $P_r f \not\to f$ uniformly on $\partial\mathbb{D}$.

11B. Fourier Series and L^p on the Unit Circle

(11.30) Orthonormal basis of $L^2(\partial\mathbb{D})$

The family $\{z^n\}_{n\in\mathbb{Z}}$ is an orthonormal basis of $L^2(\partial\mathbb{D})$.

Proof sketch: Let $H = \overline{\operatorname{span}}\{z^n : n \in \mathbb{Z}\} \subset L^2(\partial\mathbb{D})$. If $f \perp H$, then $\widehat{f}(n) = \langle f, z^n\rangle = 0$ for all n. Given $\varepsilon > 0$, choose $g \in C^2(\partial\mathbb{D})$ with $\|f - g\|_2 < \varepsilon$ (density). For $g \in C^2$, one has $|\widehat{g}(n)| \lesssim 1/n^2$, hence its Fourier series converges absolutely and uniformly, so $g \in H$ and therefore $\langle f, g\rangle = 0$. Then

$$\|f\|_2^2 = \langle f, f\rangle = \langle f, f - g\rangle \leq \|f\|_2\,\|f - g\|_2 < \varepsilon\|f\|_2,$$

so $\|f\|_2 < \varepsilon$; since ε is arbitrary, $f = 0$ and thus $H = L^2(\partial\mathbb{D})$.

(11.31) L^2-convergence of Fourier series

If $f \in L^2(\partial\mathbb{D})$, then

$$f = \sum_{n\in\mathbb{Z}} \widehat{f}(n)\,z^n \quad \text{with convergence in } L^2(\partial\mathbb{D}).$$

(11.32) Example: computing $\sum_{n\geq 1}\frac{1}{n^2}$

Let $f(e^{it}) = t$ for $t \in (-\pi, \pi]$ (extended 2π-periodically). Then $\widehat{f}(0) = 0$, and for $n \neq 0$,

$$\widehat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} t\,e^{-int}\,dt = \frac{i(-1)^n}{n}.$$

Hence

$$\sum_{n\in\mathbb{Z}} |\widehat{f}(n)|^2 = 2\sum_{n=1}^{\infty}\frac{1}{n^2}, \qquad \|f\|_2^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} t^2\,dt = \frac{\pi^2}{3}.$$

By Parseval,

$$\frac{\pi^2}{3} = 2\sum_{n=1}^{\infty}\frac{1}{n^2} \quad \Rightarrow \quad \sum_{n=1}^{\infty}\frac{1}{n^2} = \frac{\pi^2}{6}.$$

(11.36) Definition: convolution on $\partial\mathbb{D}$

For $f, g \in L^1(\partial\mathbb{D})$, define

$$(f * g)(z) = \int_{\partial\mathbb{D}} f(w)\, g(z\overline{w})\, d\sigma(w),$$

for those z for which the integral is well-defined (in fact for a.e. z).

(11.37) Convolution preserves L^1

If $f, g \in L^1(\partial\mathbb{D})$, then $f * g$ is defined a.e., $f * g \in L^1(\partial\mathbb{D})$, and

$$\|f * g\|_1 \le \|f\|_1\, \|g\|_1.$$

(11.38) Young's inequality (one form)

If $1 \le p \le \infty$, $f \in L^p(\partial\mathbb{D})$ and $g \in L^1(\partial\mathbb{D})$, then

$$\|f * g\|_p \le \|f\|_p\, \|g\|_1.$$

(11.41) Convolution is commutative

If $f, g \in L^1(\partial\mathbb{D})$, then $f * g = g * f$.

(11.42) Approximate identity (Poisson kernel)

For $1 \le p < \infty$ and $f \in L^p(\partial\mathbb{D})$,

$$\|P_r * f - f\|_p \longrightarrow 0 \qquad (r \to 1^-).$$

(11.43) Uniqueness in L^1

If $f \in L^1(\partial\mathbb{D})$ and $\widehat{f}(n) = 0$ for all $n \in \mathbb{Z}$, then $f = 0$ a.e.

(11.44) Fourier coefficients of a convolution

If $f, g \in L^1(\partial\mathbb{D})$, then

$$\widehat{(f * g)}(n) = \widehat{f}(n)\, \widehat{g}(n) \qquad \forall n \in \mathbb{Z}.$$

(11.46) Convolution is associative

If $f, g, h \in L^1(\partial\mathbb{D})$, then

$$(f * g) * h = f * (g * h).$$

Exercise (1): a real ONB on $[-\pi, \pi]$

Define $(e_k)_{k\in\mathbb{Z}}$ on $[-\pi, \pi]$ by

$$e_k(t) = \begin{cases} \frac{1}{\sqrt{\pi}} \sin(kt), & k > 0, \\ \frac{1}{\sqrt{2\pi}}, & k = 0, \\ \frac{1}{\sqrt{\pi}} \cos(|k|t), & k < 0. \end{cases}$$

Then $\{e_k\}_{k\in\mathbb{Z}}$ is an orthonormal basis of $L^2([-\pi, \pi])$. (Normalization follows from $\int_{-\pi}^{\pi} \sin^2(nt)\, dt = \int_{-\pi}^{\pi} \cos^2(nt)\, dt = \pi$, and orthogonality from standard trig product-to-sum identities.)

Exercise (3): computing $\sum_{n\geq 1} \frac{1}{n^4}$

Let $f(x) = x^2$ on $(-\pi, \pi)$ (extended 2π-periodically). Then f is even, so $b_n = 0$ and

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} x^2\,dx = \frac{2\pi^2}{3}, \qquad a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x^2\cos(nx)\,dx = \frac{4(-1)^n}{n^2} \quad (n \geq 1).$$

Parseval gives

$$\frac{1}{\pi}\int_{-\pi}^{\pi} x^4\,dx = \frac{a_0^2}{2} + \sum_{n=1}^{\infty} a_n^2.$$

Compute

$$\frac{1}{\pi}\int_{-\pi}^{\pi} x^4\,dx = \frac{2}{\pi}\int_{0}^{\pi} x^4\,dx = \frac{2\pi^4}{5}, \quad \frac{a_0^2}{2} = \frac{1}{2}\left(\frac{2\pi^2}{3}\right)^2 = \frac{2\pi^4}{9}, \quad \sum_{n\geq 1} a_n^2 = 16\sum_{n\geq 1}\frac{1}{n^4}.$$

Thus

$$\frac{2\pi^4}{5} = \frac{2\pi^4}{9} + 16\sum_{n=1}^{\infty}\frac{1}{n^4} \quad \Rightarrow \quad \sum_{n=1}^{\infty}\frac{1}{n^4} = \frac{\pi^4}{90}.$$

Exercise (6): real-valued functions and Fourier symmetry

For $f \in L^1(\partial\mathbb{D})$, f is real-valued a.e. iff

$$\widehat{f}(-n) = \overline{\widehat{f}(n)} \qquad \forall n \in \mathbb{Z}.$$

Exercise (7): L^2 iff square-summable Fourier coefficients

If $f \in L^1(\partial\mathbb{D})$, then

$$f \in L^2(\partial\mathbb{D}) \quad \Longleftrightarrow \quad \sum_{n\in\mathbb{Z}} |\widehat{f}(n)|^2 < \infty.$$

($\Rightarrow$ is Parseval. For $\Leftarrow$, use Riesz–Fischer to build $g \in L^2$ with $\widehat{g}(n) = \widehat{f}(n)$, note $g \in L^1$ since $\partial\mathbb{D}$ has finite measure, and conclude $f = g$ a.e. by L^1 uniqueness.)

Exercise (8): unimodular functions and an autocorrelation identity

Let $f \in L^2(\partial\mathbb{D})$. Then $|f(z)| = 1$ a.e. on $\partial\mathbb{D}$ iff for every $n \in \mathbb{Z}$,

$$\sum_{k\in\mathbb{Z}} \widehat{f}(k)\,\overline{\widehat{f}(k-n)} = \begin{cases} 1, & n = 0, \\ 0, & n \neq 0. \end{cases}$$

Idea: set $h(z) = |f(z)|^2 - 1 = f(z)\overline{f(z)} - 1 \in L^1$ (by Hölder). Then $h = 0$ a.e. iff $\widehat{h}(n) = 0$ for all n. Using the Fourier-coefficient formula for products (cf. Exercise (15) below) gives

$$\widehat{f\overline{f}}\,(n) = \sum_{k\in\mathbb{Z}} \widehat{f}(k)\,\overline{\widehat{f}(k-n)},$$

and $\widehat{\mathbf{1}}(n) = \delta_{n0}$, so $\widehat{h}(n) = 0$ is exactly the displayed condition.

Exercise (10): convolution operator on L^2

Let $f \in L^1(\partial\mathbb{D})$ and define $T : L^2(\partial\mathbb{D}) \to L^2(\partial\mathbb{D})$ by $Tg = f * g$.

(a) T is compact on $L^2(\partial\mathbb{D})$. Indeed, with $e_n(z) = z^n$ (an ONB), (11.44) gives

$$T(e_n) = f * e_n = \widehat{f}(n)\, e_n,$$

so T is diagonal with eigenvalues $\lambda_n = \widehat{f}(n) \to 0$ by Riemann–Lebesgue; hence T is compact.

(b) T is injective iff $\widehat{f}(n) \neq 0$ for all $n \in \mathbb{Z}$. (If $\widehat{f}(n_0) = 0$ then $T(e_{n_0}) = 0$; conversely, if all $\widehat{f}(n) \neq 0$ then $Tg = 0$ forces $\widehat{g}(n) = 0$ for all n.)

(d) The adjoint satisfies $T^* = T_{f^*}$ where

$$f^*(z) = \overline{f(\overline{z})} \qquad (z \in \partial\mathbb{D}).$$

Equivalently, $\widehat{f^*}(n) = \overline{\widehat{f}(n)}$, so $T^*(e_n) = \overline{\widehat{f}(n)}\, e_n$.

Exercise (11): convolution in angular coordinates

If $f, g \in L^1(\partial\mathbb{D})$ and $z = e^{it}$, then

$$(f * g)(e^{it}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{ix})\, g\Big(e^{i(t-x)}\Big)\, dx,$$

obtained by parametrizing $w = e^{ix}$ in $(f * g)(z) = \int f(w) g(z\overline{w})\, d\sigma(w)$.

Exercise (15): Fourier coefficients of a product

If $f, g \in L^2(\partial\mathbb{D})$, then for each $n \in \mathbb{Z}$,

$$\widehat{(fg)}(n) = \sum_{k \in \mathbb{Z}} \widehat{f}(k)\, \widehat{g}(n-k),$$

and the series converges absolutely by Cauchy–Schwarz since $(\widehat{f}(k))_{k\in\mathbb{Z}}, (\widehat{g}(k))_{k\in\mathbb{Z}} \in \ell^2(\mathbb{Z})$.

Exercise (18): Wirtinger's inequality

If $f : \mathbb{R} \to \mathbb{R}$ is C^1, 2π-periodic, and $\int_{-\pi}^{\pi} f(t)\, dt = 0$, then

$$\int_{-\pi}^{\pi} f(t)^2\, dt \ \leq\ \int_{-\pi}^{\pi} (f'(t))^2\, dt,$$

with equality iff $f(t) = a \sin t + b \cos t$ for some constants a, b.

Proof outline via Fourier series: write

$$f(t) = \sum_{n=1}^{\infty} \big(a_n \cos(nt) + b_n \sin(nt)\big) \quad (a_0 = 0 \text{ by the mean-zero condition}),$$

so Parseval yields $\int f^2 = \pi \sum_{n\geq 1} (a_n^2 + b_n^2)$. Differentiate termwise,

$$f'(t) = \sum_{n=1}^{\infty} n\big(-a_n \sin(nt) + b_n \cos(nt)\big),$$

and Parseval gives $\int (f')^2 = \pi \sum_{n\geq 1} n^2 (a_n^2 + b_n^2) \geq \pi \sum_{n \geq 1} (a_n^2 + b_n^2) = \int f^2$. Equality forces $a_n = b_n = 0$ for all $n \geq 2$, hence $f(t) = a_1 \cos t + b_1 \sin t$.

Chapter IIC: Fourier Transform

Conventions

Throughout this section, $L^p(\mathbb{R}) = L^p(\lambda)$ where λ is Lebesgue measure on $\mathbb{R}$, and for $1 \le p < \infty$,

$$\|f\|_p = \left(\int_{\mathbb{R}} |f(x)|^p \, dx \right)^{1/p}, \qquad \|f\|_\infty = \operatorname*{ess\,sup}_{x \in \mathbb{R}} |f(x)|.$$

(II.47) Definition: Fourier transform

For $f \in L^1(\mathbb{R})$, the **Fourier transform** of f is the function $\widehat{f} : \mathbb{R} \to \mathbb{C}$ defined by

$$\widehat{f}(t) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i t x} \, dx.$$

(*Remark:* Some texts use $\mathcal{F}_0 f(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-i\xi x} \, dx$, which differs from the above by a rescaling of frequency.)

Example (Fourier transform of $e^{-2\pi|x|}$). If $f(x) = e^{-2\pi|x|}$, then for $t \in \mathbb{R}$,

$$\widehat{f}(t) = \int_{-\infty}^{\infty} e^{-2\pi|x|} e^{-2\pi i t x} \, dx = \int_0^{\infty} e^{-2\pi x} e^{-2\pi i t x} \, dx + \int_{-\infty}^{0} e^{2\pi x} e^{-2\pi i t x} \, dx = \frac{1}{\pi(1+t^2)}.$$

(II.48) Result: Riemann–Lebesgue lemma

If $f \in L^1(\mathbb{R})$, then $\widehat{f}$ is uniformly continuous on $\mathbb{R}$ and

$$\|\widehat{f}\|_\infty \le \|f\|_1, \qquad \lim_{|t| \to \infty} \widehat{f}(t) = 0.$$

(II.50) Result: Derivative of a Fourier transform

Suppose $f \in L^1(\mathbb{R})$ and define $g(x) = x f(x)$. If $g \in L^1(\mathbb{R})$, then $\widehat{f}$ is continuously differentiable and

$$(\widehat{f})'(t) = -2\pi i \, \widehat{g}(t) = -2\pi i \, \widehat{(xf)}(t), \qquad t \in \mathbb{R}.$$

Example: $e^{-\pi x^2}$ equals its Fourier transform. Let $f(x) = e^{-\pi x^2}$. Using integration by parts (or the identity above),

$$(\widehat{f})'(t) = -2\pi i \int_{\mathbb{R}} x e^{-\pi x^2} e^{-2\pi i t x} \, dx = -2\pi t \, \widehat{f}(t).$$

But $f'(t) = -2\pi t f(t)$ as well, hence $\left(\frac{\widehat{f}}{f} \right)' = 0$, so $\widehat{f} = cf$ for some constant c. Evaluating at $t = 0$ gives

$$c = \widehat{f}(0) = \int_{\mathbb{R}} e^{-\pi x^2} \, dx = 1,$$

so $\widehat{f} = f$.

(II.54) Result: Fourier transform of a derivative

Suppose $f \in L^1(\mathbb{R})$ is continuously differentiable and $f' \in L^1(\mathbb{R})$. (Under these hypotheses one also has $\lim_{|x|\to\infty} f(x) = 0$.) Then, for $t \in \mathbb{R}$,

$$\widehat{f'}(t) = 2\pi i t\, \widehat{f}(t).$$

(II.56) Result: Translations, modulations, and dilations

Let $f \in L^1(\mathbb{R})$ and $b \in \mathbb{R}$.

- If $g(x) = f(x-b)$, then $\widehat{g}(t) = e^{-2\pi i t b}\, \widehat{f}(t)$.
- If $g(x) = e^{2\pi i b x} f(x)$, then $\widehat{g}(t) = \widehat{f}(t-b)$.
- If $b \neq 0$ and $g(x) = f(bx)$, then $\widehat{g}(t) = \dfrac{1}{|b|}\, \widehat{f}\left(\dfrac{t}{b}\right)$.

Example (Fourier transform of a "rotated" exponential)

Fix $y > 0$ and $x \in \mathbb{R}$, and define

$$h(t) = e^{-2\pi y|t|}\, e^{2\pi i x t}.$$

Let $f_0(t) = e^{-2\pi|t|}$ so that $\widehat{f_0}(s) = \dfrac{1}{\pi(1+s^2)}$. Since $e^{-2\pi y|t|} = f_0(yt)$, by dilation,

$$(\widehat{e^{-2\pi y|t|}})(s) = \frac{1}{y}\widehat{f_0}\left(\frac{s}{y}\right) = \frac{y}{\pi(s^2+y^2)}.$$

Then, by modulation,

$$\widehat{h}(s) = \frac{y}{\pi\big((s-x)^2+y^2\big)}.$$

(II.58) Result: Integral of a function times a Fourier transform

If $f, g \in L^1(\mathbb{R})$, then

$$\int_{\mathbb{R}} \widehat{f}(t)\, g(t)\, dt = \int_{\mathbb{R}} f(t)\, \widehat{g}(t)\, dt.$$

(Proof: expand $\widehat{f}(t)$ and apply Fubini/Tonelli.)

Consequence (Poisson kernel identity). With $h(t) = e^{-2\pi y|t|}e^{2\pi i x t}$ as above, for $f \in L^1(\mathbb{R})$,

$$\int_{\mathbb{R}} \widehat{f}(t)\, e^{-2\pi y|t|}e^{2\pi i x t}\, dt = \int_{\mathbb{R}} f(t)\, \widehat{h}(t)\, dt = \frac{1}{\pi}\int_{\mathbb{R}} f(t)\, \frac{y}{(x-t)^2+y^2}\, dt.$$

(II.63) Definition: Convolution on $\mathbb{R}$

Let $f, g : \mathbb{R} \to \mathbb{C}$ be measurable. The **convolution** $f * g$ is defined by

$$(f*g)(x) = \int_{-\infty}^{\infty} f(t)\, g(x-t)\, dt$$

for those x for which the integral makes sense.

(II.64) Result: L^p-convolution estimate (Young's inequality, special case)

If $1 \le p \le \infty$, $f \in L^1(\mathbb{R})$, and $g \in L^p(\mathbb{R})$, then $f * g$ is defined for a.e. x and

$$\|f * g\|_p \le \|f\|_1 \, \|g\|_p.$$

(*Note:* On $\mathbb{R}$, neither $L^1(\mathbb{R})$ nor $L^p(\mathbb{R})$ is contained in the other in general.)

(II.65) Result: Commutativity

If f, g are measurable and $(f * g)(x)$ is defined, then $(f * g)(x) = (g * f)(x)$.

(II.66) Result: Fourier transform of a convolution

If $f, g \in L^1(\mathbb{R})$, then

$$\widehat{(f * g)}(t) = \widehat{f}(t)\, \widehat{g}(t), \qquad t \in \mathbb{R}.$$

(II.67) Definition: Upper half-plane

Let

$$\mathbb{H} = \{z = x + iy \in \mathbb{C} : y > 0\}$$

be the upper half-plane, and identify its boundary with the real line $\partial\mathbb{H} \cong \mathbb{R}$.

(II.68) Definition: Poisson kernel

For $y > 0$, define $P_y : \mathbb{R} \to (0, \infty)$ by

$$P_y(x) = \frac{1}{\pi} \frac{y}{x^2 + y^2}.$$

The family $\{P_y\}_{y>0}$ is the **Poisson kernel** on $\mathbb{H}$. Key properties:

- $P_y(x) > 0$ for all $x \in \mathbb{R}$, $y > 0$.
- $\displaystyle\int_{\mathbb{R}} P_y(x)\, dx = 1$ for all $y > 0$.
- For every $\delta > 0$, $\displaystyle\int_{|x| \ge \delta} P_y(x)\, dx \to 0$ as $y \downarrow 0$.

(II.70) Definition: Poisson integral $P_y f$

For $f \in L^p(\mathbb{R})$ (some $p \in [1, \infty]$) and $y > 0$, define

$$(P_y f)(x) = (f * P_y)(x) = \int_{\mathbb{R}} f(t)\, P_y(x - t)\, dt = \frac{1}{\pi} \int_{\mathbb{R}} f(t)\, \frac{y}{(x-t)^2 + y^2}\, dt.$$

(II.71) Result: Uniform approximation (bounded uniformly continuous case)

If $f : \mathbb{R} \to \mathbb{C}$ is bounded and uniformly continuous, then

$$\lim_{y \downarrow 0} \|f - P_y f\|_\infty = 0,$$

i.e. $P_y f \to f$ uniformly on $\mathbb{R}$ as $y \downarrow 0$.

(II.72) Result: Poisson integral is harmonic

Let $f \in L^p(\mathbb{R})$ for some $p \in [1, \infty)$. Define $u : \mathbb{H} \to \mathbb{C}$ by

$$u(x, y) = (P_y f)(x), \qquad x \in \mathbb{R},\ y > 0.$$

Then u is harmonic on $\mathbb{H}$.

Proof sketch. Write $z = x + iy$ and note

$$\frac{y}{(x-t)^2 + y^2} = -\Im\left(\frac{1}{z-t}\right).$$

Hence

$$u(x, y) = -\Im\left(\frac{1}{\pi}\int_{\mathbb{R}} \frac{f(t)}{z-t}\, dt\right).$$

The integral defines an analytic function of $z \in \mathbb{H}$ (justified by Hölder and dominated convergence), so its imaginary part is harmonic.

(II.73) Result: Dirichlet problem on the upper half-plane

Suppose $f : \mathbb{R} \to \mathbb{C}$ is bounded and uniformly continuous. Define $u : \overline{\mathbb{H}} \to \mathbb{C}$ by

$$u(x + iy) = \begin{cases} (P_y f)(x), & y > 0, \\ f(x), & y = 0. \end{cases}$$

Then u is continuous on $\overline{\mathbb{H}}$, harmonic on $\mathbb{H}$, and $u|_{\partial\mathbb{H}} = f$.

(II.74) Result: L^p convergence

If $1 \le p < \infty$ and $f \in L^p(\mathbb{R})$, then

$$\lim_{y \downarrow 0} \|f - P_y f\|_p = 0.$$

(One standard proof uses Minkowski/Hölder and the function $h(t) = \int_{\mathbb{R}} |f(x) - f(x+t)|^p\, dx$, observing h is bounded, uniformly continuous, and $h(0) = 0$, then bounding $\|f - P_y f\|_p^p$ by $(P_y h)(0)$.)

(II.76) Result: Fourier inversion formula

If $f \in L^1(\mathbb{R})$ and $\widehat{f} \in L^1(\mathbb{R})$, then for a.e. $x \in \mathbb{R}$,

$$f(x) = \int_{\mathbb{R}} \widehat{f}(t)\, e^{2\pi i t x}\, dt.$$

Equivalently, $\widehat{\widehat{f}}(x) = f(-x)$ (a.e.), hence $f(x) = \widehat{\widehat{f}}(-x)$ (a.e.).

(II.80) Result: Functions are determined by their Fourier transforms

If $f \in L^1(\mathbb{R})$ and $\widehat{f}(t) = 0$ for all $t \in \mathbb{R}$, then $f = 0$ a.e.

(II.81) Result: Convolution is associative

If $f, g, h \in L^1(\mathbb{R})$, then
$$(f * g) * h = f * (g * h) \qquad \text{(a.e.)}.$$

(II.82) Result: Plancherel theorem

If $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, then
$$\|\widehat{f}\|_2 = \|f\|_2.$$
In particular, the map $f \mapsto \widehat{f}$ extends uniquely to a unitary operator on $L^2(\mathbb{R})$.

(II.85) Definition: Fourier transform on $L^2(\mathbb{R})$

The Fourier transform $\mathcal{F}$ on $L^2(\mathbb{R})$ is the bounded operator on $L^2(\mathbb{R})$ such that
$$\mathcal{F}f = \widehat{f} \qquad \text{for all } f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}).$$

(II.87) Properties of $\mathcal{F}$ on $L^2(\mathbb{R})$

- $\mathcal{F}$ is unitary on $L^2(\mathbb{R})$.
- $\mathcal{F}^4 = I$ (in fact $\mathcal{F}^2 f(x) = f(-x)$).
- $\sigma(\mathcal{F}) \subseteq \{1, i, -1, -i\}$, and in fact $\sigma(\mathcal{F}) = \{1, i, -1, -i\}$.

Exercises IIC

(1) Let $f \in L^1(\mathbb{R})$. Prove that $\|\widehat{f}\|_\infty = \|f\|_1$ *iff* there exist $\theta \in \mathbb{R}$ and $t_0 \in \mathbb{R}$ such that
$$e^{i\theta} f(x) e^{-2\pi i t_0 x} \geq 0 \quad \text{for a.e. } x \in \mathbb{R}.$$
(Equivalently: after multiplying by a constant phase and a character, f is a.e. nonnegative real.)

(2) Suppose $f(x) = xe^{-\pi x^2}$ for all $x \in \mathbb{R}$. Show that
$$\widehat{f} = -if.$$
(*Hint:* Use $\widehat{e^{-\pi x^2}} = e^{-\pi t^2}$ and the identity $\widehat{(xg)} = \frac{i}{2\pi}(\widehat{g})'$.)

(3) Suppose
$$f(x) = 4\pi x^2 e^{-\pi x^2} - e^{-\pi x^2} = (4\pi x^2 - 1)e^{-\pi x^2}.$$
Show that
$$\widehat{f} = -f.$$

(6) Suppose
$$f(x) = \begin{cases} xe^{-2\pi x}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$
Show that for all $t \in \mathbb{R}$,
$$\widehat{f}(t) = \frac{1}{4\pi^2(1+it)^2}.$$
(Compute $\widehat{f}(t) = \int_0^\infty xe^{-2\pi(1+it)x}\,dx$.)

(8) Let $f \in L^1(\mathbb{R})$ and $n \in \mathbb{Z}_{>0}$. Define $g : \mathbb{R} \to \mathbb{C}$ by $g(x) = x^n f(x)$. Prove that if $g \in L^1(\mathbb{R})$, then $\widehat{f}$ is n-times continuously differentiable on $\mathbb{R}$ and

$$(\widehat{f})^{(n)}(t) = (-2\pi i)^n \widehat{g}(t) = (-2\pi i)^n \widehat{(x^n f)}(t), \qquad t \in \mathbb{R}.$$

Chapter 12: Probability Measures

(12.1) Definition: Probability measure

Let $\mathcal{F}$ be a σ-algebra on a set Ω. A *probability measure* on $(\Omega, \mathcal{F})$ is a measure P such that

$$P(\Omega) = 1.$$

Here Ω is the *sample space*. An *event* is a set $A \in \mathcal{F}$, and $P(A)$ is the probability of A. If P is a probability measure on $(\Omega, \mathcal{F})$, then $(\Omega, \mathcal{F}, P)$ is a *probability space*.

(12.3) Definition: Indicator function

If $A \subseteq \Omega$, the indicator (characteristic) function of A is

$$\mathbf{1}_A : \Omega \to \mathbb{R}, \qquad \mathbf{1}_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A. \end{cases}$$

(12.4) Definition: Almost surely

Let $(\Omega, \mathcal{F}, P)$ be a probability space. An event A happens *almost surely* (a.s.) if

$$P(A) = 1 \quad \Longleftrightarrow \quad P(\Omega \setminus A) = 0.$$

Example (informal): "Picking a random real number gives an irrational number a.s."

(12.6) Result: Borel–Cantelli Lemma (I)

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(A_n)_{n\geq 1}$ a sequence of events with

$$\sum_{n=1}^{\infty} P(A_n) < \infty.$$

Then

$$P(A_n \text{ i.o.}) = 0, \qquad \text{where} \quad \{A_n \text{ i.o.}\} = \limsup_{n\to\infty} A_n = \bigcap_{n\geq 1} \bigcup_{m\geq n} A_m.$$

Proof sketch. Let $A = \limsup A_n$. For each n,

$$\mathbf{1}_A \leq \sum_{m\geq n} \mathbf{1}_{A_m}.$$

Integrate and use monotone convergence to get

$$P(A) = E[\mathbf{1}_A] \leq \sum_{m\geq n} P(A_m) \xrightarrow[n\to\infty]{} 0.$$

(12.7) Definition: Independent events

Events $A, B \in \mathcal{F}$ are *independent* if

$$P(A \cap B) = P(A)P(B).$$

A family $\{A_k\}_{k \in I} \subseteq \mathcal{F}$ is *independent* if for any distinct $k_1, \dots, k_n \in I$,

$$P\Big(\bigcap_{j=1}^{n} A_{k_j}\Big) = \prod_{j=1}^{n} P(A_{k_j}).$$

(12.9) Example: Independence in a product space

If $(\Omega_1, \mathcal{F}_1, P_1)$ and $(\Omega_2, \mathcal{F}_2, P_2)$ are probability spaces, then

$$(\Omega_1 \times \Omega_2,\ \mathcal{F}_1 \otimes \mathcal{F}_2,\ P_1 \times P_2)$$

is a probability space. If $A \in \mathcal{F}_1$ and $B \in \mathcal{F}_2$, then the events

$$A \times \Omega_2 \quad \text{and} \quad \Omega_1 \times B$$

are independent since

$$(A \times \Omega_2) \cap (\Omega_1 \times B) = A \times B, \qquad (P_1 \times P_2)(A \times B) = P_1(A)P_2(B).$$

(12.10) Result: Borel–Cantelli Lemma (II) (Reverse, under independence)

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(A_n)_{n \geq 1}$ an independent family of events such that

$$\sum_{n=1}^{\infty} P(A_n) = \infty.$$

Then

$$P(A_n \text{ i.o.}) = 1.$$

Proof sketch. Let $A = \limsup A_n$. Then $A^c = \bigcup_{n \geq 1} \bigcap_{m \geq n} A_m^c$. By independence,

$$P\Big(\bigcap_{m \geq n} A_m^c\Big) = \prod_{m \geq n} \big(1 - P(A_m)\big).$$

Using $\log(1-x) \leq -x$ for $x \in (0,1)$, this product is $\leq \exp\big(-\sum_{m \geq n} P(A_m)\big) \to 0$, hence $P(A^c) = 0$.

(12.13) Definition: Random variable; expectation

A *random variable* on $(\Omega, \mathcal{F}, P)$ is a measurable function $X : \Omega \to \mathbb{R}$. If $X \in L^1(P)$, its *expectation* is

$$E[X] = \int_{\Omega} X \, dP.$$

Since $P(\Omega) = 1$, $E[X]$ can be interpreted as the average/mean of X.

(12.14) Definition: Independent random variables

Random variables X, Y are *independent* if for all Borel sets $U, V \subseteq \mathbb{R}$, the events $\{X \in U\}$ and $\{Y \in V\}$ are independent. More generally, $\{X_k\}_{k \in I}$ is independent if for any distinct $k_1, \ldots, k_n$ and Borel sets $U_1, \ldots, U_n$,

$$P\Big(\bigcap_{j=1}^{n} \{X_{k_j} \in U_j\}\Big) = \prod_{j=1}^{n} P(X_{k_j} \in U_j).$$

(12.16) Result: Functions of independent r.v.'s are independent

If X, Y are independent and $f, g : \mathbb{R} \to \mathbb{R}$ are Borel measurable, then $f(X)$ and $g(Y)$ are independent. **Proof sketch.** For Borel U, V,

$$\{f(X) \in U\} = \{X \in f^{-1}(U)\}, \qquad \{g(Y) \in V\} = \{Y \in g^{-1}(V)\}.$$

(12.17) Result: Expectation of the product of independent r.v.'s

If X, Y are independent and integrable (e.g. $X, Y \in L^1(P)$ and $XY \in L^1(P)$), then

$$E[XY] = E[X]\,E[Y].$$

Proof sketch. Verify first for simple functions, then extend by approximation and standard convergence arguments.

(12.19) Definition: Variance and standard deviation

If $X \in L^2(P)$, define

$$\mathrm{var}(X) = E\big[(X - E[X])^2\big], \qquad \sigma(X) = \sqrt{\mathrm{var}(X)}.$$

Example: for an event A,

$$\mathrm{var}(\mathbf{1}_A) = E[(\mathbf{1}_A - P(A))^2] = P(A)\big(1 - P(A)\big), \quad \text{so} \quad \sigma(\mathbf{1}_A) = \sqrt{P(A)(1 - P(A))}.$$

(12.20) Result: Variance formula

If $X \in L^2(P)$, then

$$\mathrm{var}(X) = E[X^2] - \big(E[X]\big)^2.$$

Proof sketch. Expand $E[(X - E[X])^2] = E[X^2] - 2E[X]E[X] + (E[X])^2$.

(12.21) Result: Chebyshev's inequality

If $X \in L^2(P)$ and $t > 0$, then

$$P\big(|X - E[X]| \geq t\,\sigma(X)\big) \leq \frac{1}{t^2}.$$

Proof sketch. Apply Markov to $(X - E[X])^2$:

$$P\big((X - E[X])^2 \geq t^2\sigma^2(X)\big) \leq \frac{E[(X - E[X])^2]}{t^2\sigma^2(X)} = \frac{1}{t^2}.$$

(12.22) Result: Variance of a sum of independent r.v.'s

If $X_1, \dots, X_n \in L^2(P)$ are independent, then

$$\operatorname{var}\Big(\sum_{k=1}^{n} X_k\Big) = \sum_{k=1}^{n} \operatorname{var}(X_k).$$

Proof sketch. Use $\operatorname{var}(S) = E[S^2] - (E[S])^2$ with $S = \sum X_k$ and independence to get $E[X_i X_j] = E[X_i]E[X_j]$ for $i \neq j$.

(12.23) Definition: Conditional probability

If $B \in \mathcal{F}$ with $P(B) > 0$, define $P(\cdot \mid B) : \mathcal{F} \to [0,1]$ by

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

(12.24) Result: Bayes' theorem (V1)

If $A, B \in \mathcal{F}$ with $P(A) > 0$ and $P(B) > 0$, then

$$P(B \mid A) = \frac{P(A \mid B)\, P(B)}{P(A)}.$$

(12.25) Result: Bayes' theorem (V2)

Let B be an event with $P(B) > 0$, and let $A_1, \dots, A_n$ be pairwise disjoint events with $P(A_i) > 0$ and $\bigcup_{i=1}^{n} A_i = \Omega$. Then for $k = 1, \dots, n$,

$$P(A_k \mid B) = \frac{P(B \mid A_k)\, P(A_k)}{\sum_{i=1}^{n} P(B \mid A_i)\, P(A_i)}.$$

(12.27) Definition: Distribution and distribution function

Let X be a random variable. Its *(probability) distribution* is the probability measure P_X on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ defined by

$$P_X(B) = P(X \in B) = P\big(X^{-1}(B)\big), \qquad B \in \mathcal{B}(\mathbb{R}).$$

Its *distribution function* (CDF) is $F_X : \mathbb{R} \to [0,1]$ given by

$$F_X(s) = P(X \leq s) = P_X((-\infty, s]).$$

(12.29) Result: Characterization of distribution functions

A function $H : \mathbb{R} \to [0,1]$ is a distribution function (i.e. $H = F_X$ for some r.v. X) iff:

1. H is nondecreasing: $s \leq t \Rightarrow H(s) \leq H(t)$;
2. $\lim_{s \to -\infty} H(s) = 0$;
3. $\lim_{s \to \infty} H(s) = 1$;
4. H is right-continuous: $\lim_{s \downarrow t} H(s) = H(t)$ for all $t \in \mathbb{R}$.

(12.32) Definition: Density function

A random variable X has a *density* (PDF) h if there exists $h \in L^1(\mathbb{R})$ with $h \geq 0$ a.e. such that for all $s \in \mathbb{R}$,

$$P(X \leq s) = \int_{-\infty}^{s} h(u)\, du.$$

(Equivalently, $P_X(B) = \int_B h(u)\, du$ for all Borel B.)

(12.33) Result: Mean and variance from a density

If $h \geq 0$ and $\int_{\mathbb{R}} h(x)\, dx = 1$, define a probability measure P on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ by

$$P(B) = \int_B h(x)\, dx.$$

Let $X : \mathbb{R} \to \mathbb{R}$ be $X(x) = x$. Then h is a density for X, and (when integrable)

$$E[X] = \int_{-\infty}^{\infty} x\, h(x)\, dx, \qquad \operatorname{var}(X) = \int_{-\infty}^{\infty} x^2 h(x)\, dx - \Big(\int_{-\infty}^{\infty} x\, h(x)\, dx \Big)^2.$$

(12.35) Definition: Identically distributed; i.i.d.

A family of random variables is *identically distributed* if all have the same distribution function. More explicitly, $\{X_k\}_{k \in I}$ is identically distributed if for all $s \in \mathbb{R}$ and all $j, k \in I$,

$$P(X_j \leq s) = P(X_k \leq s).$$

If additionally the family is independent, it is *i.i.d.* (When moments exist, identically distributed implies $E[X_j] = E[X_k]$ and $\sigma(X_j) = \sigma(X_k)$.)

(12.38) Result: Weak Law of Large Numbers

Let $\{X_k\}_{k \geq 1}$ be an i.i.d. family with $X_k \in L^2(P)$ and $E[X_k] = \mu$. Then for every $\varepsilon > 0$,

$$\lim_{n \to \infty} P\Big(\Big| \frac{1}{n} \sum_{k=1}^{n} X_k - \mu \Big| \geq \varepsilon \Big) = 0.$$

Proof sketch. Let $\overline{X}_n = \frac{1}{n} \sum_{k=1}^{n} X_k$. Then $E[\overline{X}_n] = \mu$ and $\operatorname{var}(\overline{X}_n) = \operatorname{var}(X_1)/n$ by independence. Chebyshev gives

$$P(|\overline{X}_n - \mu| \geq \varepsilon) \leq \frac{\operatorname{var}(X_1)}{n\varepsilon^2} \xrightarrow[n \to \infty]{} 0.$$