Analysis Qualifying Exam Study Guide

Stepan Malkov

Last Updated: January 2026

The following is a collection of notes and exercises intended to help prepare for the UCLA Analysis Qualifying Exam. The content is split into four major areas: Real Analysis, Harmonic Analysis, Functional Analysis, and Complex Analysis. Due to the high variability of topics appearing on the exam, there is an emphasis on content breadth. Additionally, there is a focus on frequently appearing exam problems and techniques. There may be a number of typos, which generally should not affect the correctness of the arguments herein.

Contents

1 Real Analysis

1.1 Distribution Theory

Definition 1.1.1. A **distribution** is a continuous linear functional on a nice function space of **test functions.** For example,

(a) The space of **distributions** $\mathcal{D}'$ is the dual of C_c^∞.

(b) The space of **tempered distributions** $\mathcal{S}'$ is the dual of the Schwartz space $\mathcal{S}$.

(c) The space of **compactly supported distributions** $\mathcal{E}'$ is the dual of C^∞.

Note that one has the inclusions

$$C_c^\infty \subset \mathcal{S} \subset C^\infty \subset \mathcal{E}' \subset \mathcal{S}' \subset \mathcal{D}',$$

where one identifies f with $(f, \cdot)_{L^2}$. We endow distribution spaces with the weak-* topology, i.e. $L_n \to L$ in $\mathcal{S}'$ if $L_n(f) \to L(f)$ for all $f \in \mathcal{S}$, and write $L(f) = \langle f, L \rangle$.

The reason for dealing with distributions is to extend the classical theory of analysis to more general objects than functions. Motivated by integration by parts, we define the **derivative** of a distribution L via the adjoint relation

$$\langle f, \partial_{x_i} L \rangle = -\langle \partial_{x_i} f, L \rangle.$$

For example, if H is the Heaviside step-function,

$$\langle f, H' \rangle = -\langle f', H \rangle = -\int_0^\infty f' dx = f(0),$$

so $H' = \delta$ is the **delta function,** defined by $\langle f, \delta \rangle = f(0)$. Similarly, the **Fourier transform** of a distribution is given by

$$\langle f, \widehat{L} \rangle = \langle \widehat{f}, L \rangle,$$

and as an example,

$$\langle f, \widehat{\delta^{(k)}} \rangle = \langle \widehat{f}, \delta^{(k)} \rangle = (-1)^k \widehat{f}^{(k)}(0) = \widehat{(-ix)^k f}(0) = \int_{-\infty}^\infty (-ix)^k f(x) dx = \langle f, (ix)^k \rangle.$$

where the latter is the Kronecker delta. in the sense of functions. A distribution's **support** S is the smallest closed set such that the distribution vanishes on any function supported outside S, giving a precise definition to $\mathcal{E}'$. One may also define the **convolution** of a distribution and a test function by

$$\langle \phi, u * T \rangle = \langle \widetilde{u} * \phi, T \rangle,$$

where $\widetilde{u}(x) = u(-x)$. More generally, if S, T are distributions and T has compact support, then $S * T$ is defined as the unique distribution satisfying $(S * T) * u = S * (T * u)$ for all test functions u. The convolution of two distributions is a commutative and associative operation satisfying the usual formulas.

1.2 Convergence of Functions

A very common type of problem in analysis is to show that a sequence of functions converges (or does not converge). Here is an outline of the main proof techniques and types of convergence:

1.2.1 Almost Everywhere Convergence

Almost Everywhere (a.e.) pointwise convergence: A sequence of (Lebesgue) measurable functions f_n is said to converge **pointwise a.e.** if

$$\mu(\{x : f_n(x) \nrightarrow f(x)\}) = 0.$$

Here are some ways to prove pointwise convergence:

Proposition 1.2.1. *If $f_n \to f$ in L^p or in measure, there is a subsequence f_{n_k} such that $f_{n_k} \to f$ pointwise a.e.*

Proof. Convergence in L^p implies convergence in measure, as

$$\int |f - f_n|^p = \int_{|f-f_n|<\epsilon} |f - f_n|^p + \int_{|f-f_n|\geqslant\epsilon} |f - f_n|^p \geqslant \epsilon^p \mu(\{\{x : |f - f_n| > \epsilon\}),$$

and taking $n \to \infty$ yields the claim. Thus, it suffices to assume that $f_n \to f$ in measure. Indeed, if for any $\epsilon > 0$, $\mu(\{x : |f_n - f| > \epsilon\}) \to 0$, one may find a subsequence n_k of sets

$$A_k = \left\{x : |f_{n_k}(x) - f(x)| > \frac{1}{2^k}\right\}, \quad \mu(A_k) < \frac{1}{2^k}.$$

By Borel-Cantelli, we get that

$$\mu(\limsup_{k\to\infty} A_k) = \mu\left(\left\{x : |f_{n_k}(x) - f(x)| > \frac{1}{2^k} \text{ for infinitely many k}\right\}\right) = 0.$$

But the negation of that statement is precisely $|f_{n_k}(x) - f(x)| \leqslant \frac{1}{2^k}$ for large enough k, i.e. $f_{n_k}(x) \to f(x)$. Thus,

$$\mu(\{x : f_n(x) \nrightarrow f(x)\} = 0,$$

so $f_{n_k} \to f$ a.e. □

Remark 1.2.1. A very useful criterion for convergence is as follows: $a_n \to a$ in a metric space iff every subsequence of $(a_n)_n$ has a further subsequence converging to a. In particular, since this is not true for a.e. convergence of the typewriter sequence, one concludes that a.e. convergence is not metrizable.

1.2.2 L^p convergence

Definition 1.2.1. We say that $f_n \to f$ **in** L^p for $1 \leqslant p < \infty$ if

$$\|f_n - f\|_p^p = \int |f_n - f|^p dx \to 0,$$

and $f_n \to f$ in L^∞ if $f_n \to f$ uniformly except on a null set.

There are three main tools and a variety of corollaries that may be used to establish convergence in L^p.

Theorem 1.2.1. *(Monotone Convergence Theorem) If $f_n \geqslant 0$ is an increasing sequence of functions and $f_n \to f$ pointwise a.e., then $f_n \to f$ in L^p.*

Proof. Let $f_n \to f$ be an increasing sequence of functions. Then, clearly, $\lim_{n\to\infty} \int f_n \leqslant \int f$. For the converse, pick a simple function g such that $\int(f - g) < \epsilon$, and consider the set of points $E_n = \{x : f_n(x) \geqslant \alpha g(x)\}$. Then,

$$\int f_n \geqslant \alpha \int_{E_n} g,$$

and as $\alpha \to 1$ and $n \to \infty$, taking the supremum over all simple g yields the claim.

The L^p case then easily follows. □

Theorem 1.2.2. *(Fatou's Lemma) If $f_n \geqslant 0$, then $\int \liminf_{n\to\infty} f_n \leqslant \liminf_{n\to\infty} \int f_n$.*

Corollary 1.2.1. *(Reverse Fatou) By applying Fatou to $g - f_n$, one gets that if $f_n \in L^1$ and $f_n \leqslant g$, $g \in L^1$, then*

$$\limsup_{n\to\infty} \int f_n \leqslant \int \limsup_{n\to\infty} f_n.$$

Theorem 1.2.3. *(Dominated Convergence Theorem) If $f_n \to f$ a.e. and $|f_n| \leqslant g$ for $g \in L^p$, then $f_n \to f$ in L^p.*

There are two particularly strong results that are necessary to prove more interesting claims regarding convergence in L^p.

Theorem 1.2.4. *(Egorov's Theorem) Let $\mu(X) < \infty$, and suppose $f_n \to f$ a.e. on X. Then, for every $\epsilon > 0$, there exists a set A such that $\mu(A) < \epsilon$ and $f_n \to f$ uniformly on A^c.*

Proof. Consider the set

$$A_{n,k} := \left\{x : |f_n(x) - f(x)| > \frac{1}{k}\right\}.$$

Recall from the proof of convergence in measure implies a.e. convergence that $\mu(\limsup_{n\to\infty} A_{n,k}) = 0$ for all k. Pick a subsequence n_k s.t.

$$B_k := \bigcap_{n \geqslant n_k} A_{n,k}$$

and $\mu(B_k) < \frac{\epsilon}{2^k}$ for all k. Then, if $C = \bigcup_k B_k, \mu(C) < \epsilon$, on C^c one has

$$\forall k \exists n_k \forall n \geqslant n_k, |f_n(x) - f(x)| \leqslant \frac{1}{k},$$

which implies uniform convergence.

□

Theorem 1.2.5. *(Lusin's Theorem) If $\mu(X) < \infty$ and f is measurable on X, for any $\epsilon > 0$ there exists a compact set K such that $\mu(K^c) < \epsilon$ and f is continuous on K.*

Proof. Take a sequence $f_n \in C_c(X)$ such that $f_n \to f$ in a.e. (which is possible by density of C_c in L^1). By Egorov and regularity of the Lebesgue measure, $f_n \to f$ uniformly on some compact set K such that $\mu(K^c) < \epsilon$, so f is the uniform limit of continuous functions on a compact set, i.e. it is continuous. □

We now list the proofs of the fundamental inequalities used in analysis.

Lemma 1.2.1 (Young's Inequality). *For $a, b \geqslant 0, \frac{1}{p} + \frac{1}{q} = 1$,*

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q},$$

with equality iff $a^p = b^q$.

Proof. Consider x^{p-1} and its inverse x^{q-1} on $[0, a] \times [0, b]$. Then, the sum of the integrals of the two functions is at least the area of the rectangle, with equality iff the functions touch the corner, i.e. $a^{p-1} = b$, so $a^p = b^q$, and the claim follows. ☐

Definition 1.2.2. If p, q satisfy the condition above, they are known as **Hölder conjugates.**

Theorem 1.2.6 (Hölder's inequality). *For $f \in L^p, q \in L^q$, where $1 \leqslant p, q \leqslant \infty$ are Hölder conjugates, then*

$$\int |fg| \leqslant \|f\|_p \|g\|_q,$$

with equality for $1 < p < \infty$ iff $|f|^p = c|g|^q$ for some $c \geqslant 0$.

Proof. Normalize f, g so that $\|f\|_p = \|g\|_q = 1$, and apply Young's inequality. The cases $p = 1, q = 1$ may be checked directly. The equality follows from the equality case in Young's inequality. ☐

Corollary 1.2.2 (Generalized Hölder). *If $\sum_{k=1}^n \frac{1}{p_k} = \frac{1}{r}$, then*

$$\left\| \prod_{k=1}^n f_k \right\|_r \leqslant \prod_{k=1}^n \|f_k\|_{p_k}.$$

Theorem 1.2.7 (Minkowski's Inequality). *For $1 \leqslant p \leqslant \infty$,*

$$\|f + g\|_p \leqslant \|f\|_p + \|g\|_p,$$

with equality for $1 < p < \infty$ iff $f = cg$ for some constant c.

Proof. Normalizing so that $\|f + g\|_p = 1$, by Hölder,

$$\int |f + g|^p \leqslant \int |f + g|^{p-1}|f| + |f + g|^{p-1}|g| \leqslant \|f + g\|_p^{p-1}(\|f\|_p + \|g\|_p) = \|f\|_p + \|g\|_p,$$

and so equality holds iff $|f + g|^{\frac{p-1}{q}} = c|f|^p = c'|g|^p$, and since $(p - 1)q = p$ and g and f must have the same sign at each point, equality holds iff $g = cf$.

☐

Theorem 1.2.8 (Dual of L^p). *For $1 \leqslant p < \infty$, the dual of L^p is L^q, where $\frac{1}{p} + \frac{1}{q} = 1$.*

Proof. For $\phi \in (L^p)^*$, define the (signed) measure $\nu : A \to \phi(\chi_A)$. Then, the desired function is the Radon-Nikodym derivative $g_\phi = \frac{d\nu}{d\mu}$, and setting $f_n = |g|^{\frac{q}{p}} \chi_{|g| \leqslant n}$ and using monotone convergence lemma yields $g \in L^q$. ☐

Definition 1.2.3. For a σ-finite measure/topological space (X, μ), define the Banach spaces

$$\text{rca}(\mu) \subset \text{ca}(\mu) \subset \text{ba}(\mu)$$

of **bounded regular Borel, countably additive, and finitely additive signed measures** absolutely continuous with respect to μ with the total variation norm.

Remark 1.2.2. The above proof shows that $(L^\infty(X,\mu))^* \cong \text{ba}(\mu)$, and $(L^1)^{**} \cong \text{ca} \subset (L^\infty)^*$.

Lemma 1.2.2. *L^1 is weakly sequentially complete, i.e. every weakly Cauchy sequence converges.*

Proof. Clearly, f_n is uniformly bounded in L^1 by Banach-Steinhaus. Define the signed measure

$$\nu(A) = \lim_{n\to\infty} \int_A f_n d\mu,$$

and consider the Radon-Nikodym derivative $f = \frac{d\nu}{d\mu}$. Then, one easily verifies that $\int f\chi_A d\mu = \nu(A) = \lim_{n\to\infty} \int f_n \chi_A d\mu$, and so $f_n \rightharpoonup f$. □

Remark 1.2.3. Since the unit ball in a reflexive Banach space is weakly sequentially compact, all reflexive Banach spaces are weakly sequentially complete.

Corollary 1.2.3. *If $\lim_{n\to\infty} \int_A f_n$ exists and is finite for all measurable A, then $f_n \rightharpoonup f$ for some $f \in L^1$.*

The following are important consequences of Egorov's Theorem as they relate to convergence in L^p :

Lemma 1.2.3. *If $\mu(X) < \infty$, $\|f_n\|_p \leqslant M < \infty$, $1 < p < \infty$, and $f_n \to f$ a.e., then $f_n \to f$ in L^1.*

Proof. First note that $f \in L^p$, as by Fatou,

$$\|f\|_p = \||f|^p\|_1 = \int \liminf_{n\to\infty} |f_n|^p \leqslant \liminf_{n\to\infty} \|f_n\|^p < \infty.$$

Fix $\epsilon > 0$. By Egorov, $f_n \to f$ uniformly on some set A, so by Hölder,

$$\int_X |f-f_n| = \int_A |f-f_n| + \int_{A^c} |f-f_n| \leqslant \epsilon\mu(X) + \|\chi_{A^c}\|_q \|f-f_n\|_p = \epsilon\mu(X) + \mu(A^c)^{\frac{1}{q}} M < \epsilon(\mu(X)+M)$$

for p, q Hölder conjugates and large enough n by choosing A such that $\mu(A^c)^{\frac{1}{q}} < \epsilon$. □

Remark 1.2.4. The same argument shows that if $\mu(X) < \infty$, $f_n \to f$ a.e. and $\|f_n\|_q \leqslant M$ for $q > p$, then $f_n \to f$ in L^p.

Lemma 1.2.4. *For $1 \leqslant p < \infty$, if $f_n \in L^p$, $f_n \to f$ a.e., and $\|f_n\|_p \to \|f\|_p$, then $f_n \to f$ in L^p.*

Proof. Note that since $|x|^p$ is convex, $|f - f_n|^p \leqslant 2^{p-1}(|f|^p + |f_n|^p)$. Applying Fatou's lemma to $2^{p-1}(|f|^p + |f_n|^p) - |f - f_n|^p$ yields

$$2^p \|f\|^p \leqslant \liminf_{n\to\infty} 2^{p-1}(\|f\|_p^p + \|f_n\|_p^p) - \limsup_{n\to\infty} \|f - f_n\|_p^p,$$

which then yields the desired inequality. □

Corollary 1.2.4. *If $f_n \to f$ a.e. and $f_n \in L^2$, $f \in L^2$ and $\|f\|_2 \leqslant \liminf_{n\to\infty} \|f_n\|_2$.*

Proof. Fatou's lemma applied to f_n^2. □

1.2.3 Uniform Integrability and Compactness in L^p

Definition 1.2.4. A subset $X \subset L^p$ is called **uniformly integrable** if it is uniformly bounded in L^p and for any $\epsilon > 0$ there exists a $\delta > 0$ such that for any $f \in X$, $\|f\|_{L^p(E)} < \epsilon$ whenever $\mu(E) < \delta$.

Remark 1.2.5. If A is a finite measure space, then X is uniformly integrable in L^1 iff for all $\epsilon > 0$, there is a λ such that $\sup_X \int_{|f|>\lambda} f(x)d\mu < \epsilon$.

Proof. Suppose X is uniformly integrable. Then, by Chebyshev, $\mu(\{x : |f| > \lambda\}) \leqslant \frac{M}{\lambda}$ for $M = \sup_X \|f\|_1$, so by uniform integrability, one can make $\int_{|f|>\lambda} f(x)d\mu < \epsilon$ for large enough λ. Conversely, we get that $\sup_X \|f\|_1 \leqslant \epsilon + \lambda\mu(A)$, and for any $\epsilon > 0$, $\int_E |f| \leqslant \epsilon + \lambda\mu(E) < 2\epsilon$ for $\delta = \frac{\epsilon}{\lambda}$. □

A lot of results regarding convergence work only on finite measure spaces, yet lots of time one works with infinite measure spaces. A very useful concept known as tightness allows us to reduce the problem to a finite measure space.

Definition 1.2.5. A family $X \subset L^p(X)$ is **tight** if for all $\epsilon > 0$, there exists $E \subset X$, $\mu(E) < \infty$, s.t. $\|f\|_{L^p(E^c)} < \epsilon$ for all $f \in X$.

Theorem 1.2.9 (Vitali Convergence Theorem)**.** *If $\{f_n\}$ is a tight sequence of uniformly integrable functions in $L^p, 1 \leqslant p < \infty$, then $f_n \to f$ in L^p iff $f_n \to f$ in measure.*

Proof. Suppose $f_n \to f$ in measure. Pick $\epsilon > 0$ and choose a corresponding E. Then, by uniform integrability, pick $\delta > 0$ such that $\|f\|_{L^p(E)} < \epsilon$ when $\mu(E) < \delta$. Moreover, by Egorov, pick $A_\epsilon \subset E$ such that $\mu(A_\epsilon^c) < \delta$. Passing to an a.e. convergent subsequence, we use Fatou's lemma to conclude that $\|f\|_{L^p(E^c)}, \|f\|_{L^p(A_\epsilon^c)} < \epsilon = (\mu(E) + 3)\epsilon$.

$$\begin{aligned}\int |f_n - f|^p dx &= \int_{E \cap A_\epsilon} |f_n - f|^p dx + \int_{E \cap A_\epsilon^c} |f_n - f|^p dx + 2\epsilon \\ &\leqslant \mu(E)\epsilon + 2\epsilon + 2\epsilon.\end{aligned}$$

□

Remark 1.2.6. Since Egorov shows that convergence a.e. on a finite measure set implies convergence in measure, one obtains the following strong corollary: if f_n is uniformly integrable in L^p, tight, and $f_n \to f$ a.e., then $f_n \to f$ in L^p.

Lemma 1.2.5. *If $f_n \in L^1$ and $\int_A f_n$ converges and is finite for all measurable A, then f_n is uniformly integrable in L^1. In particular, if $f_n \to f$ in L^p, then $\{f_n\}$ is uniformly integrable.*

A very important theorem is that of precompactness in L^p spaces.

Theorem 1.2.10 (Kolmogorov-Riesz)**.** *$X \subset L^p(\Omega), \Omega \subset \mathbb{R}^n, 1 \leqslant p < \infty$ is precompact iff:*

(a) ***Boundedness:*** $\sup_{f \in X} \|f\|_p < \infty$.

(b) ***Tightness:*** *For any $\epsilon > 0$, there exists an $R > 0$ such that $\int_{[-R,R]^c} |f|^p < \epsilon$ for all $f \in X$.*

(c) ***(Uniform) Continuity:*** *For all $\epsilon > 0$, there exists a $\delta > 0$ such that $\int |f(x+y) - f(x)|^p dx < \epsilon$ whenever $|y| < \delta$ for all $f \in X$.*

Remark 1.2.7. Tightness and uniform continuity imply boundedness, so it is not strictly necessary.

Remark 1.2.8. If Ω is bounded, uniform continuity is the only required condition of the theorem.

Proof. ($\Longrightarrow$): Suppose X satisfies the conditions of the theorem, and fix ϵ, R, y as in the statement. Let Q be an open cube centered at the origin, let Q_i be nonoverlapping translates such that their closure covers $B(x,R)$, and define Pf on Q_i to be the average of f on Q_i and 0 otherwise. Then, by Minkowski and noting that $x, y \in Q_i$ implies $x - y \in 2Q$,

$$\begin{aligned}\|f - Pf\|_p^p &< \epsilon^p + \sum_i \int_{Q_i} \left| \frac{1}{\mu(Q_i)} \int_{Q_i} (f(x) - f(z)) dz \right|^p dx \\ &\leqslant \epsilon^p + \sum_i \int_{Q_i} \frac{1}{\mu(Q)} \int_{2Q} |f(x) - f(x+y)|^p dy dx \\ &\leqslant \epsilon^p + \frac{1}{\mu(Q)} \int_{2Q} \int_{\mathbb{R}^n} |f(x) - f(x+y)|^p dx dy \leqslant (2^n + 1)\epsilon^p.\end{aligned}$$

□

($\Longleftarrow$): If $X \subset L^p(\Omega)$ is precompact, it is clearly bounded. Moreover, for $\epsilon > 0$, if $B(f_1, \epsilon), .., B(f_n, \epsilon)$, cover X, pick R s.t. $\|f_i\|_{L^p(B(x,R))} < \epsilon$. Then,

$$\|f\|_{L^p(B(x,R)^c)} \leqslant \|f - f_i\|_{L^p(B(x,R)^c)} + \|f_i\|_{L^p(B(x,R)^c)} < 2\epsilon.$$

Uniform continuity is established almost exactly the same way.

1.2.4 Weak and Weak-* Convergence in L^p

Definition 1.2.6. For $1 \leqslant p < \infty$, let q be the Hölder conjugate of p. We say that $f_n \rightharpoonup f$ (or f_n **converges weakly to** f) in L^p if $\langle f_n, g \rangle \to \langle f, g \rangle$ for all $g \in L^q$, where the inner product is just $\langle f, g \rangle = \int f\overline{g}$. For $1 < p \leqslant \infty$, we say $f_n \overset{*}{\rightharpoonup} f$ (or f_n **converges weak-* to** f) in L^p if $\langle f_n, g \rangle \to \langle f, g \rangle$ for all $g \in L^q$. Note that weak and weak-* convergence are the same for $1 < p < \infty$.

There are two examples to keep in mind when proving weak convergence:

Example 1.2.1 (The traveling wave)**.** If $f \in L^p, 1 < p < \infty$ then if $f_n(x) = f(x+n), f \rightharpoonup 0$ in L^p.

Proof. For simplicity, we consider the case when $f \in L^p(\mathbb{R})$. For any $g \in L^q$, Pick compact sets A, B such that $\|f\|_{L^p(A^c)}, \|g\|_{L^q(B^c)} < \epsilon$. Then, for $n > \sup_{x \in A, y \in B} |x - y|$, $B \cap (A - n) = \varnothing$, where $A - n = \{a - n : a \in A\}$. Thus, by Hölder,

$$\begin{aligned}|\langle f_n, g \rangle| &= \left| \int f_n \overline{g} \right| \\ &\leqslant \int |f(x+n)\overline{g}(x)| \\ &\leqslant \int_B |f(x+n)\overline{g}(x)| + \int_{A-n} |f(x+n)\overline{g}(x)| + \int_{(B \cup (A-n))^c} |f(x+n)\overline{g}(x)| \\ &\leqslant \epsilon\|g\|_q + \epsilon\|f\|_p + \epsilon^2,\end{aligned}$$

□

where $\|f\|_{L^p(B)} \leqslant \epsilon$ since $(B + n) \cap A = \varnothing$. Sending $\epsilon \to 0$ completes the proof.

Remark 1.2.9. Clearly, if one chooses $g = 1$, this need not hold for $p = 1$.

Example 1.2.2 (The Oscillator)**.** Let $f \in L^\infty$ be a k-periodic function. Then, if $f_n(x) = f(nx)$, $f_n \xrightarrow{*} \frac{1}{k}\int_0^k f(x)dx$ in L^1.

Corollary 1.2.5. *Suppose $f_n \rightharpoonup f$ in L^p. Then, $\|f_n\|_p$ is uniformly bounded, and $\|f\|_p \leqslant \liminf \|f_n\|_p$.*

Proof. Weakly convergent sequence are bounded + weak lower-semicontinuity of the norm. ☐

Clearly, the examples show that weak convergence need not imply convergence a.e., convergence in measure, or L^p convergence. Conversely, it is easy to see using Hölder's that convergence in L^p implies weak convergence. The vertical blow-up examples shows that convergence in measure or a.e. convergence do not necessarily imply weak convergence. The following important theorem provides a criterion for weak convergence:

Theorem 1.2.11 (Dunford-Pettis Theorem)**.** *Let $X \subset L^1$. Then, X is uniformly integrable iff it weakly precompact.*

Proof. If X is uniformly integrable, the weak-* closure $\overline{X}^* \subset (L^\infty)^*$ is weak-* compact by Banach-Alaoglu. A finitely additive map F is countably additive iff $\lim_n F(A_n) = 0$ for $\bigcap_n A_n = \varnothing$, where A_n is a decreasing sequence of sets. Then, uniform integrability implies that any $F \in \overline{X}^*$ is countably additive, and thus is given by integration against some $f \in L^1$, and there exists a sequence $i^{-1}(f_n) \rightharpoonup i^{-1}(f)$ in L^1. In particular, $i^{-1} : \overline{X}^* \to L^1$ is weak$-$*-weak continuous. Thus, $i^{-1}(\overline{X}^*)$ is weakly compact, so X is weakly precompact.

Conversely, suppose X is weakly precompact and not uniformly integrable, and pick a nonuniformly integrable subsequence f_n such that $\int_{|f_n|>n} |f_n| d\mu \geqslant C$ for some $C > 0$. Then, by Eberlein-Smullian, any sequence has a weakly convergent subsequence. But this implies that the subsequence is uniformly integrable (see Lemma 2.9), a contradiction. ☐

1.2.5 Exercises

Problem 1.2.1 (Spring 2010 Problem 1). Show that a sequence that converges in L^p has an a.e. convergent subsequence. Moreover, find a sequence of functions that converges to 0 in L^2 that does not converge a.e.

Proof. The first part is Proposition 1.1. For the second part, one may take the typewriter sequence $f_1 = \chi_{[0,1]}, f_2 = \chi_{[0,\frac{1}{2}]}, \chi_3 = \chi_{[\frac{1}{2},1]}, \ldots$ which converges to 0 in L^2 but not a.e. ☐

Problem 1.2.2 (Spring 2012 Problem 1). Let $1 < p < \infty$, $f_n : \mathbb{R}^3 \to \mathbb{R}$ such that $\limsup \|f_n\|_p < \infty$. Show that if $f_n \to f$ a.e, then $f_n \to f$ weakly.

Proof. For any $g \in L^q$, pick a compact set A such that $\|g\|_{L^q(A^c)} < \epsilon$ and $\|g\|_{L^q(E)} < \epsilon$ whenever $\mu(E) < \delta$ for a small enough $\delta > 0$. Then, by Egorov's theorem,

$$\left|\int (f_n - f)\overline{g}\right| \leqslant \int_{A\setminus E} |(f_n - f)\overline{g}| + \int_E |(f_n - f)\overline{g}| + \int_{A^c} |(f_n - f)\overline{g}| = O(\epsilon),$$

where the first term is bounded by uniform convergence on a compact set, second term is bounded since $\mu(E) < \delta$, and the third term is bounded by choosing A to be large enough. ☐

Problem 1.2.3 (Spring 2014 Problem 3). Suppose $f_n \to 0$ a.e., $\|f_n\|_2 < \infty$. Show $f_n \rightharpoonup 0$ in L^2.

Proof. This is a specific case of the previous problem. □

Problem 1.2.4 (September 2018 Problem 1). Suppose $f_n \to f$ a.e., $\sup_n \|f_n\|_2 < \infty$, and $\sup_n \|xf_n\|_1 < \infty$. Show that $f_n, f \in L^1$, $f_n \to f$ in L^1, and that neither of the last two conditions may be omitted.

Proof. The second condition implies that on $[-M,M]^c$, $\|f_n\|_1 \leqslant \frac{N}{M}$ for some fixed $N > 0$. In particular, on $[-M,M]$, lemma 1.6 guarantees that $f_n \to f \in L^1$. Thus, for any $\epsilon > 0$, pick M and n large enough so that $\int_{[-M,M]^c} |f| < \epsilon$, so then

$$\int |f_n - f| = \int_{[-M,M]} |f_n - f| + \int_{[-M,M]^c} |f_n - f| \leqslant \frac{\epsilon}{2} + (\frac{N}{M} + \epsilon) < 2\epsilon$$

for sufficiently large M.
Neither of the last two conditions may be omitted, as demonstrated by the counterexamples $f_n = \chi_{[n,n+1]}$ and $g_n = n^2\chi_{[0,\frac{1}{n}]}$. □

Problem 1.2.5 (Fall 2020 Problem 2). Show that there exists a constant c such that

$$\langle f, \cos(\sin(n\pi x))\rangle \to \langle f, c\rangle$$

for all $f \in L^1$.

Proof. Note that $\cos(\sin(n\pi x))$ is 2-periodic. Thus, $c = \frac{1}{2}\int_0^2 \cos(\sin(n\pi x))dx$ by the oscillator example. □

Problem 1.2.6 (Fall 2010 Problem 3). Let $f_n(x) = e^{\sin(2\pi nx)}$. Show f_n converges weakly in $L^1([0,1])$ and weak-* in $L^\infty([0,1])$.

Proof. Let $f(x) = e^{\sin(2\pi x)}$, and note that f is 1-periodic. Then, by the weak convergence lemma, $f_n \overset{*}{\rightharpoonup} \int_0^1 e^{\sin(2\pi x)}dx$ in L^∞. Moreover, f_n is uniformly bounded in $L^\infty([0,1])$. By density arguments, to show $f_n \rightharpoonup \int_0^1 e^{\sin(2\pi x)}dx$ in L^1, it suffices again to consider characteristic functions of closed intervals. But this is indeed already shown by the weak convergence in L^∞ argument, so the proof is complete. □

Problem 1.2.7 (Spring 2020 Problem 2). Let f_n be a sequence of differentiable functions satisfying $\sup \|f_n\|_1 < \infty, \sup \|f_n'\|_1 < \infty$, and for any $\epsilon > 0$, there exists an $R(\epsilon) > 0$ such that $\sup \|f_n\|_{L^1([-R,R]^c)} < \epsilon$. Show f_n is has a convergent subsequence in L^1.

Proof. We use Riesz-Fischer. Clearly, the first two conditions establish uniform boundedness and tightness. The third condition and Minkowski shows that for $|y| < \delta$,

$$\int |f(x+y) - f(x)| = \int \int_x^{x+y} |f'(t)| dt dx \leqslant \int_x^{x+y} \|f'\|_1 \leqslant |y| M$$

by the uniform bound on the derivatives. Thus, the conditions for Riesz-Fischer are satisfied, showing that $\{f_n\}$ has a convergent subsequence in L^1. □

Problem 1.2.8 (Spring 2017 Problem 2). Let $f_n : [0,1] \to [0,\infty)$ be a sequence of nondecreasing functions uniformly bounded in L^2. Show that there exists a subsequence that converges in L^1.

Proof. We apply Riesz-Fischer. Indeed, since f_n is uniformly bounded in L^2, it is uniformly bounded in L^1. Since the sequence is on a finite measure space, tightness is unnecessary. Finally, to show continuity, we use the fact that f_n is nondecreasing. Namely, this implies that each f_n has at most a countable number of discontinuities, so each f_n agrees with a continuous function a.e.
Then, for any $\epsilon > 0$,

$$\int_0^{1-y} |f(x+y) - f(x)| dx$$

□

Problem 1.2.9. Show that $l^1(\mathbb{N})$ has the **Schur property,** i.e. every weakly convergent sequence is norm-convergent.

Proof. Suppose $x_n \rightharpoonup x$ but $x_n \not\to x$ in l^1. Then, there exists a subsequence satisfying $\|x_{n_k} - x\| \geqslant \epsilon$. By a diagonalization argument, pick a further subsequence where the ith element in each subsequence has the same sign for all i. Now, let S_k be a finite subset of the support of $x_{n_k} - x$ s.t. $\|x_{n_k} - x\|_{l^1(S_k)} \geqslant \frac{\epsilon}{2}$. Note that $\bigcup_k S_k$ cannot be bounded, as otherwise $x_n \not\to x$ on a finite set. In particular, norm convergence on finite sets implies that after passing to another subsequence, there exist a sequence of pairwise disjoint $A_i \subseteq S_i$ s.t. $\|x_{n_k} - x\|_{l^1(A_k)} \geqslant \frac{\epsilon}{4}$. Then, letting $y = \operatorname{sign}(x_{n_1}) 1_{\bigcup_k A_k}$, we see that

$$(x_{n_k} - x) \cdot y \geqslant \frac{\epsilon}{4},$$

for all k, which is a contradiction. Thus, $x_n \to x$ in l^1. □

Theorem 1.2.12 (Conclusions). *(a) If f_n converges a.e. and is bounded by an integrable function, apply the Dominated Convergence Theorem to get convergence in L^p.*

(b) If $f_n \to f$ converges a.e. and $\|f_n\|_p \to \|f\|_p$, then $f_n \to f$ in L^p.

(c) If $\mu(X) < \infty$, $\sup \|f_n\|_p < \infty$ and $f_n \to f$ a.e., then $f_n \to f$ in L^q for $q < p$.

(d) If $\mu(X) < \infty$ and $f_n \to f$ a.e., $f_n \to f$. Additionally, if f_n is uniformly integrable, $f_n \to f$ in L^p.

(e) If $\sup \|f_n\|_p < \infty$ *and* $f_n \to f$ *a.e.,* $f_n \rightharpoonup f$.

(f) X *is precompact in* L^p *iff it is tight, continuous, and uniformly bounded.*

1.3 Lebesgue Differentiation Theorem

Here we cover an extremely important theorem that allows us to "differentiate" L^p functions. For this, we first need to introduce a lot of heavy machinery.

Definition 1.3.1. For $f \in L^1_{loc}$, define the **Hardy-Littlewood maximal function** to be

$$H(f)(x) = \sup_{x \in B} \frac{1}{|B|} \int_B |f| dt,$$

where B is an open ball. In other words, H is the maximal average of f on any ball containing x.

We begin with some basic properties.

Proposition 1.3.1. *$H(f)$ is measurable and finite a.e. Moreover,* $Hf : L^1 \to L^{1,w}$ *is of weak type* $(1, 1)$.

Proof. Note that $\{x : H(f) > \lambda\}$ are open, since if $\sup_{x \in B} \frac{1}{|B|} \int_B |f| dt > \lambda$, for all points y nearby, $y \in B$, and so the supremum is also greater than λ. We now introduce with a key lemma.

Lemma 1.3.1 (Vitali Covering Lemma)**.** *Given a cover by open balls of a metric space X, there exists a finite subset* $B_{n_1}, ..., B_{n_k}$ *such that* $3B_{n_1}, ..., 3B_{n_k}$ *is a cover of X.*

Proof. Inductively pick balls of the largest radius disjoint from all the ones currently picked, and let $Y = 3B_{n_1} \cup ... \cup 3B_{n_k}$. If B is one of the balls picked, then $B \subset Y$. Otherwise, by maximality B intersects at least one of these balls B_k, and so $B \subset 3B_j$. □

Now, if $E_\lambda = \{x : H(f) > \lambda\}$, then for each E_λ is covered by open balls B_n with

$$\frac{1}{\lambda} \int_{B_n} |f| dt > |B_n|,$$

so covering a compact subset $K \subset E_\lambda$ by finitely many balls using the lemma, one obtains

$$|K| \leqslant 3^d \sum_{i=1}^{k} |B_{n_k}| \leqslant \frac{3^d}{\lambda} \int_{\mathbb{R}^d} |f| dt,$$

where we use the fact that the balls are disjoint in the integral over $\mathbb{R}^d$. Since the Lebesgue measure is regular, we are done. Moreover, the weak bound implies that $\mu\{f^* = \infty\} = 0$, so f^* is finite a.e. □

Corollary 1.3.1. *Since $H(f)$ is trivially of strong type* (∞, ∞), *by the Marcinkiewicz interpolation theorem, $H(f)$ is of strong type* (p, p) *for* $1 < p \leqslant \infty$.

What follows is a powerful consequence known as the Lebesgue differentiation theorem.

Theorem 1.3.1 (Lebesgue Differentiation Theorem)**.** *If $f \in L^1_{loc}$, for a.e. x,*

$$\lim_{x\in B, |B|\to 0} \int_B |f(y) - f(x)| dy = 0.$$

In particular,

$$\lim_{x\in B, |B|\to 0} \int_B f(y) dy = f(x)$$

a.e. Any x for which this holds is called a ***Lebesgue point,*** *implying that a.e. point is a Lebesgue point of f.*

Proof. It suffices to show that the set

$$E_\lambda = \{x : \limsup \left| \frac{1}{|B|} \int_B f(y) - f(x) dy \right| > 2\lambda\}$$

has measure 0 for all λ. Approximating f in L^1 with a continuous function g with compact support so that $\|f - g\|_1 < \epsilon$ and noting that the limsup vanishes for continuous functions,

$$\limsup \frac{1}{|B|} \left| \int_B f(y) - f(x) dy \right| \leqslant (f-g)^*(x) + |f(x) - g(x)|.$$

If

$$F_\lambda = \{x : (f-g)^*(x) > \lambda\}, G_\lambda = \{x : |f(x) - g(x)| > \lambda\},$$

then $E_\lambda \subset G_\lambda \cup F_\lambda$. But by Chebyshev and Hardy-Littlewood for $f - g$, this implies

$$|E_\alpha| \leqslant |G_\alpha| + |F_\alpha| \leqslant \frac{C}{\alpha}\epsilon,$$

and sending $\epsilon \to 0$ completes the proof. Now, for the general case, enumerate the rationals and apply the proof to the function $|f(y) - r|$, with $E = \bigcup_r E_r$, where E_r is the set where the previous theorem fails. Then, for $x \notin E$,

$$\frac{1}{|B|} \int_B |f(y) - f(x)| dy \leqslant \frac{1}{|B|} \int_B |f(y) - r| dy + |f(x) - r|,$$

and the proof is complete. □

Corollary 1.3.2. *Let $d\nu = d\lambda + f d\mu$ be the Lebesgue-Radon-Nikodym representation of ν. Then,*

$$\lim_{r\to 0} \frac{\nu(E_r)}{\mu(E_r)} = f(x).$$

for any family E_r shrinking nicely to x and a.e. x.

Proof. It suffices to prove that

$$\lim_{r\to 0} \frac{\lambda(E_r)}{\mu(E_r)} = 0$$

for a.e. x. WLOG, assume that E_r are open balls. We will show that

$$F_k = \left\{ x \in A : \limsup_{r\to 0} \frac{\lambda(B(x,r))}{\mu(B(x,r))} \right\} > \frac{1}{k}$$

has measure zero, where A contains the support of μ. Recall that since $\lambda \perp \mu$, by regularity of λ, one may pick A such that $\lambda(A) < \epsilon$. By the same argument as in the proof of Hardy-Littlewood, we cover compact subsets K of F_k by balls on which

$$\mu(K) \leqslant 3^d \sum_{k=1}^{n} \mu(B_{n_k}) \leqslant 3^d k\lambda(A) < 3^d k\epsilon,$$

and we are done. □

Example 1.3.1. There cannot exist a subset $A \subset [0,1]$ such that $\mu(A) < 1$ and $\mu(A \cap B) > \lambda\mu(B)$ for $\lambda \in (0,1)$ and all balls B, as LDT would imply that $\chi_A \geqslant \epsilon \implies \chi_A = 1$ a.e.

Corollary 1.3.3. *A monotonic function f is differentiable a.e.*

Proof. Recall that to any right-continuous, monotone function F there exists an associated Borel measure μ_F such that $\mu_F((a,b]) = F(b) - F(a)$. The Lebesgue-Radon-Nikodym derivative of this measure is (up to some minor technicalities) our derivative F'. □

1.4 BV and the Fundamental Theorem of Calculus

A fundamental result from undergraduate analysis is the Fundamental Theorem of Calculus, which states:

Theorem 1.4.1. *(FTOC)*

(a) If f is continuous, $F(x) = \int_0^x f(t)dt$ is differentiable and $F'(x) = f(x)$.

(b) If F is an antiderivative of a Riemann integrable function f, then $F(b) - F(a) = \int_a^b f(t)dt$.

Our goal in this section is to prove the most general version of this theorem.

1.4.1 Bounded Variation

Definition 1.4.1. A function is said to be of **bounded variation (BV)** on $[a,b]$ if for any sequence of intervals as above, $\sum_{i=0}^{n} |f(b_i) - f(a_i)| < \infty$. We define the **total variation function** T_F of F to be

$$T_F(x) = \sup_{b_n = x} \sum_{i=0}^{n} |f(b_i) - f(a_i)|.$$

More generally, define the class $BV(\Omega)$ of functions of **bounded variation** as a subspace of L^1 such that the **total variation**

$$V(u) := \sup_{\phi} \int_{\Omega} u \ \mathrm{div}(\phi) dx < \infty,$$

where $\|\phi\|_\infty \leqslant 1$ and ϕ is a C^1 vector field on Ω. Note that for $\phi \approx -\frac{\nabla u}{|\nabla u|}$, this gives that $V(u) \leqslant \int_\Omega |\nabla u| dx = \|\nabla u\|_1$ whenever ∇u is well-defined.

Remark 1.4.1. More simply put, $BV(\Omega)$ is the space of functions u with norm $\|u\|_{TV} = \|u\|_1 + V(u)$ whose distributional derivative Du is a finite Radon measure and satisfies

$$\langle \mathrm{div}(\phi), u \rangle = \langle \phi, Du \rangle.$$

This can be seen by defining the action of the linear functional Du according to the above formula on C^1, extending to C^0 by Hahn-Banach and constructing the appropriate measure using the Riesz Representation Theorem. Additionally, $\|Du\|_{TV} = V(u)$.

Remark 1.4.2. Note that on $W^{1,1}(\Omega) \subset BV(\Omega)$, $\|u\|_{TV} = \|u\|_{W^{1,1}}$. In general, however, functions in $W^{1,1}(\Omega)$ cannot have jump discontinuities since they admit weak derivatives, and so the space $BV(\Omega)$ is strictly larger than $W^{1,1}(\Omega)$.

Theorem 1.4.2 (Helly's Selection Theorem)**.** *Let $u_n : \mathbb{R} \to \mathbb{R}$ be a sequence of increasing functions uniformly bounded in L^p for $p > 1$. Show that u_n has a subsequence that converges in L^q_{loc} for $q < p$.*

Proof. Let K be the complement of the set of points of discontinuity of any of the functions, which is countable and therefore measure 0. Extract a convergent subsequence $u_{n_k} \to u$ on K. Then, extend u according to $u(x) = \limsup_{y \leqslant x} u(y)$. Clearly, u is positive, increasing and monotone. By regularity of Lebesgue measure, if u is continuous at x, $u_n(q_1) \leqslant u_n(x) \leqslant u_n(q_2)$ implies as $n \to \infty$ that

$$|u_n(x) - u(x)| \leqslant \sup\{|u(q_2) - u_n(q_1)|, |u_n(q_1) - u(q_2)|\},$$

and for q_1, q_2 close enough to x and n large enough, this is bounded by ϵ. Thus, $u_n \to u$ pointwise except at most on a countable set. Picking a further subsequence, one may assume that u_{n_k} converges to u pointwise everywhere. The rest of the argument follows immediately from uniform integrability and the Vitali convergence theorem. ☐

Proposition 1.4.1. *$BV(\Omega)$ is a Banach algebra with the norm $\|f\|_{BV} = \|f\|_1 + V(f)$, and V is convex lower semi-continuous on L^1 and continuous on $BV(\Omega)$.*

Proof. Lower semi-continuity of V follows from Fatou's lemma, which directly implies that $BV(\Omega)$ is Banach. We take as a given that functions in $BV(\Omega)$ satisfy the chain rule, and therefore the product rule. This implies that the product of BV functions is BV, so $BV(\Omega)$ is in fact a Banch algebra. ☐

Proposition 1.4.2. *The inclusion $BV(\Omega) \hookrightarrow L^1(\Omega)$ is compact.*

Proof. Recall Rellich-Kondrachov, which says that if Ω is bounded and p^* is the Sobolev conjugate of p, then $W^{1,p}(\Omega)$ embeds into $L^q(\Omega)$ for $1 \leqslant q \leqslant p^*$, where for $q < p^*$ the embedding is compact. Approximating a BV function u by smooth functions with uniformly bounded derivatives and applying Rellich-Kondrachov then yields a convergent subsequence in L^1. ☐

Remark 1.4.3. On $\mathbb{R}$, the compact embedding is a consequence of Helly's selection theorem, since a family of uniformly bounded monotone functions is precompact and every BV function is a sum of monotone functions.

Proposition 1.4.3. *Monotone functions are differentiable a.e. with derivative in L^1_{loc}.*

Proof. WLOG, suppose $f : [0, \infty) \to [0, \infty)$ is increasing and $f(0) = 0$. Then, f defines a premeasure according to $\mu((b - a]) = f(b) - f(a)$, which then extends to the corresponding **Lebesgue-Stiltjes measure** *df* by Caratheodory's extension theorem. Then, by the Radon-Nikodym theorem, one can write $\mu = \lambda + \rho$, where λ is absolutely continuous with respect to the Lebesgue measure m. Moreover, by the Lebesgue differentiation theorem, one has that $f' = \frac{d\lambda}{dm}$ a.e. In particular, one immediately sees that $\int_a^b f' \leqslant f(b) - f(a)$, with equality iff $\rho = 0$, which implies that $f' \in L^1_{loc}$. ☐

Remark 1.4.4. Lebesgue's differentiation theorem gives us a unique decomposition of every monotonic function $F = F_{AC} + F_d + F_s$, where F_{AC} is absolutely continuous (and therefore continuous), F_d is a jump function, and F_s is a continuous singular function with derivative 0 a.e.

It is immediately clear that (bounded) monotonic functions are of bounded variation. One now aims to obtain a decomposition of a BV function.

Proposition 1.4.4. *$F \in BV$ iff $F = \frac{1}{2}(T_F + F) - \frac{1}{2}(T_F - F)$, where $T_F + F, T_F - F$ are increasing.*

Proof. Note that $T_F(b) - T_F(a) \geqslant |F(a) - F(b)|$ by the definition of T_F. Conversely, the sum of two BV functions is still BV. ☐

Remark 1.4.5. This is known as the **Jordan decomposition** of f.

Corollary 1.4.1. *Since monotone functions are continuous except at most on a countable set, so are functions in $BV([a,b])$. In fact, BV functions are differentiable a.e. (with derivative in L^1). One may ask whether the converse is true, but it is not. Indeed, adding a bunch of Cantor functions alternating according to a conditionally convergent series convergent to 0 on intervals with rational endpoints shows that there exists a function which is differentiable a.e. with derivative in L^1, but is not BV on any subinterval.*

Remark 1.4.6. On $\mathbb{R}$, we see that BV functions are preicsely those functions whose derivatives are signed Lebesgue-Stieltjes measures.

1.4.2 Absolute Continuity and FTOC

We want to develop a generalization of the FTOC to Lebesgue measurable functions. For that, we first need to understand the properties of integrals of L^1 functions. For $f \in L^1$, $\int_a^x f(t)dt$ is easily seen to be continuous, but it is in fact in a stronger class of so-called absolutely continous functions.

Definition 1.4.2. A function f is **absolutely continuous (AC)** on $[a,b]$ if for any finite set of disjoint open intervals $(a_0, b_0), ..., (a_n, b_n)$, $a_i < b_i < a_{i+1}$, for every $\epsilon > 0$ there exists $\delta > 0$ such that $\sum_{i=0}^n |f(b_i) - f(a_i)| < \epsilon$ whenever $\sum_{i=0}^n |b_i - a_i| < \delta$.

From the definition, we immediately see that $AC \subset BV$. Moreover, note that for $F \in BV$, the corresponding signed Lebesgue-Stieltjes measure $\mu_F := \mu_+ - \mu_-$, where μ_+, μ_- are the measures corresponding to the Jordan decomposition of F, satisfies $\mu_F \ll m$ (where m is the Lebesgue measure) iff the FTOC holds (since the singular part of the Lebesgue decomposition is trivial). We now claim that this condition is precisely that of F being absolutely continuous.

Lemma 1.4.1. *$\mu_F \ll m$ iff F is absolutely continuous.*

Proof. The forward direction is immediate by applying absolutely continuity to a disjoint union of open intervals. Conversely, if $m(E) = 0$, by regularity there exist open $U_1 \supset U_2 \supset ...$ converging to E, which are a countable union of open intervals. Then $\mu_F(U_j) < \epsilon$ for large enough j in the limit of taking $N \to \infty$ intervals, so $\mu_F(E) = 0$. ☐

Corollary 1.4.2. *This argument directly shows that a continuous function F of bounded variation is absolutely continuous iff $m(E) = 0 \implies \mu_F(E) = m(F(E)) = 0$.*

To answer the question of whether the two are equal, we first describe a generalization of the classical FTOC for the Lebesgue integral.

Proposition 1.4.5. *If f is **everywhere** differentiable and $f' \in L^1$, then f is absolutely continuous, i.e. the FTOC holds.*

Proof. Find a lower semi-continuous g s.t. $g > f'$ and $\int g < \int f' + \epsilon$. Define

$$F_\eta(x) = \int_a^x g(t)dt - (f(x) - f(a)) + \eta(x - a).$$

For $t > x$ close enough to x, one can ensure that

$$g(t) > f'(x), \frac{f(t) - f(x)}{t - x} < f'(x) + \eta.$$

Then,

$$F_\eta(t) - F_\eta(x) > (t - x)f'(x) - (t - x)(f'(x) + \eta) + \eta(t - x) = 0.$$

Since F_η is continuous and $\eta > 0$ is arbitrary , this implies that $F_\eta(b) \geqslant 0$, i.e. $f(b) - f(a) \leqslant \int_a^b g < \int f' + \epsilon$, and applying the same argument to $-f$ concludes the proof. ☐

Remark 1.4.7. The Cantor function shows that differentiability everywhere cannot be weakened to differentiability a.e. However, a nontrivial generalization of this statement lets one relax everywhere differentiability to f being differentiable everywhere except on at most a countable set.

We now want to classify absolutely continuous functions in terms of BV functions. Clearly, absolutely continuous functions are continuous. WLOG, suppose that for some $F \in L^1$, $T_F(-\infty) = 0$. Then, we have the following theorem:

Theorem 1.4.3. *(FTOC, Lebesgue Version) TFAE:*

(a) F is absolutely continuous.

(b) There exists $f \in L^1$ s.t. $F(x) = \int_a^x f(t)dt$.

(c) F is differentiable a.e. with $F' \in L^1$ and $F(b) - F(a) = \int_a^b F'(t)dt$.

Proof. (b) $\implies$ (a): This follows from the fact that integrals define absolutely continuous measures w.r.t. to the Lebesgue measure.
(c) $\implies$ (b): Trivial.
(a) $\implies$ (c): Since $AC \subset BV$, F has a derivative f defined a.e. Moroever, we have shown that if F is absolutely continuous, then $\mu_F \ll m$, and so by Lebesgue decomposition, the FTOC is satisfied. ☐

To summarize, here are the properties of AC and BV functions:

(a) $F \in BV$ iff F is the sum of monotone functions iff there exists a Lebesgue-Stieltjes measure μ_F, in which case F is continuous except at most on a countable set and differentiable a.e. with $F' \in L^1$. The oscillating Cantor function shows that the converse is not true.

(b) $F \in BV$ is absolutely continuous iff $\mu_F \ll m$ iff the FTOC holds iff F is continuous and $m(E) = 0 \implies m(F(E)) = 0$, as the singular part of the Lebesgue decomposition $d\mu_f = \frac{d\mu_F}{dm} dm + d\lambda$ is zero. In particular, if F is differentiable except on at most a countable set with $F' \in L^1$, F is absolutely continuous.

Theorem 1.4.4 (The Generalized Fundamental Theorem of Calculus)**.** *As a consequence of this machinery, we have the following general characterization of the Fundamental Theorem of Calculus:*

(a) If f is differentiable everywhere, f' need not be (improper) Riemann integrable (see Volterra function) or Lebesgue integrable (if f' is unbounded but is improper Riemann integrable).

(b) If f' is (improper) Riemann integrable, then the FTOC holds (by the standard proof). If f' is Lebesgue integrable, then the FTOC holds. This is not true for a.e. differentiable functions f (see Cantor function), but is true for functions differentiable except at most on a countable set.

(c) If f is Lebesgue integrable, then by the Lebesgue differentiation theorem, $F(x) = \int_a^x f(t)dt$ is differentiable a.e. with derivative equal to f a.e. In particular, it is equal to f at x_0 iff x_0 is a Lebesgue point of f, e.g. if f is continuous at x_0. F need not be differentiable except on a countable set (see $d(x, C)$ where C is the Cantor set). Thus, if L is the set of Lebesgue points, then L is dense and L^c has measure zero (but need not be countable).

Corollary 1.4.3 (Rademacher's Theorem)**.** *Locally Lipschitz functions are differentiable a.e. and satisfy the FTOC.*

Remark 1.4.8. This theorem sounds a lot less nice than what could be easily remembered. In particular, it requires some sort of integrability for f'. One approach to this is to generalize the both the improper Riemann integral and the Lebesgue integral to the **Henstock-Kurzweil (HK) integral.** If considering HK integrals, the FTOC can then be simply phrased as: If F is differentiable with derivative f, then

$$\text{HK}\int_a^b f(x)dx = F(b) - F(a).$$

1.4.3 Hölder Continuity

Definition 1.4.3. For an open bounded subset $X \subset \mathbb{R}^n$ and $k \in \mathbb{N}, \alpha > 0$, define the **Hölder class** $C^{k,\alpha}(X)$ to be the space of C^k functions with finite Hölder norm

$$\|f\|_{C^{k,\alpha}} = \|f\|_{C^k} + |\partial^\beta f|_{C^{0,\alpha}} := \|f\|_{C^k} + \sup_{x \neq y \in X, |\beta| = k} \frac{|\partial^\beta f(x) - \partial^\beta f(y)|}{\|x - y\|^\alpha} < \infty,$$

for the Hölder seminorm $|\cdot|_{C^{0,\alpha}}$. If $f \in C^{0,\alpha}$, we say f is α-Hölder continuous.

Proposition 1.4.6. *$C^{k,\alpha}$ only contains constants for $\alpha > 1$, and $C^{k,1}$ is the vector space of k-times continuously differentiable functions with the k-th order derivative Lipschitz continuous. In particular, for $\alpha < 1$ and X bounded $C^{k,\alpha}(\overline{X})$ is a Banach space.*

Proof. The characterization of $C^{0,1}$ follows immediately from the definition. Notice that $C^{k,\alpha} \subset C^{k,\alpha'}$ whenever $\alpha > \alpha'$. WLOG suppose $k = 0$. Then,

$$\frac{|f(x) - f(y)|}{\|x - y\|} \leqslant \|f\|_{C^{0,\alpha}} \|x - y\|^{\alpha - 1} \to 0$$

as $\|x - y\| \to 0$, so ∇f is zero, i.e. f is constant. Clearly, the $C^{k,\alpha}$ norm is a norm. Finally, since $C^k(\overline{X})$ is Banach, if f_n is Cauchy in $C^{k,\alpha}$, it converges to an element $f \in C^k$. Moreover, if $|f|_{C^{0,\alpha}} > \lim_n |f_n|_{C^{0,\alpha}}$, there exists a sequence of pairs (x_k, y_k) such that

$$\frac{|f(x_k) - f(y_k)|}{\|x_k - y_k\|^\alpha} \leqslant \frac{2\epsilon}{\|x_k - y_k\|^\alpha} + |f_n|_{C^{0,\alpha}} \leqslant \limsup_n |f_n|_{C^{0,\alpha}} = \lim_n |f_n|_{C^{0,\alpha}},$$

which is a contradiction if that is the sequence of pairs that maximizes $|f|_{C^{0,\alpha}}$. Thus, $f \in C^{k,\alpha}$ and $|f|_{C^{0,\alpha}} \leqslant \lim_n |f_n|_{C^{0,\alpha}}$. Then, since $|f_n - f_m|_{C^{0,\alpha}} < \epsilon$, taking the limit in C^k yields $|f - f_m|_{C^{0,\alpha}} \leqslant \lim_n |f_n - f_m|_{C^{0,\alpha}} \leqslant \epsilon$. Thus, we conclude that $f_n \to f$ in $C^{k,\alpha}$, so $C^{k,\alpha}$ is a Banach space. □

Example 1.4.1. For $0 < \beta \leqslant 1$, $f(x) = x^\beta$ on $[0, 1]$ is α-Hölder continuous for $\alpha \leqslant \beta$, but not for $\alpha > \beta$, since

$$\sup_{x > y} \frac{x^\beta - y^\beta}{(x - y)^\beta} < \infty$$

for $\alpha \leqslant \beta$ - this can be seen since for fixed y, as $x \to y$, the function approaches 0, and x^β grows faster than $(x - y)^\beta$ since $(x - y)^{\beta - 1} \geqslant x^{\beta - 1}$. For $y = 0$ and $\alpha > \beta$, as $x \to 0$, one has $|x^\beta|_{C^{0,\alpha}} \to \infty$.

Example 1.4.2. Absolute continuity does not imply Hölder continuity, as $\frac{1}{\ln x}$ (taken to be 0 at 0) is absolutely continuous on $[0, \frac{1}{2}]$, but $\frac{\frac{1}{\ln x}}{x^\alpha} \to \infty$ for any $\alpha > 0$ as $x \to 0$.

Lemma 1.4.2. *The inclusion $C^{k,\alpha} \hookrightarrow C^{k,\beta}$ for $\alpha > \beta$ is compact.*

Proof. The inclusion is clearly continuous, as

$$|f|_{C^{0,\beta}} \leqslant \|x-y\|^{\alpha-\beta}|f|_{C^{0,\alpha}} \leqslant \operatorname{diam}(X)^{\alpha-\beta}|f|_{C^{0,\alpha}}.$$

Moreover, the sequence is uniformly equicontinuous, so by Arzela-Ascoli, there is a uniformly convergent subsequence, and

$$|f_n - f_m|_{C^{0,\beta}} \leqslant |f_n - f_m|_{C^{0,\alpha}}^{\frac{\beta}{\alpha}} \|f_n - f_m\|_\infty^{1-\frac{\beta}{\alpha}} \to 0$$

as $n, m \to \infty$ since f_n is bounded in $C^{0,\alpha}$. □

Example 1.4.3. Hölder continuous functions need not be of bounded variation. As an example, consider the **Weierstrass function**

$$f(x) = \sum_{n=1}^{\infty} a^n \cos(b^n \pi x)$$

for b odd, $0 < a < 1$, and $ab > 1 + \frac{3}{2}\pi$. By the Weierstrass M-Test, this function is continuous. Let $\alpha_n \in \mathbb{Z}$ be the integer closest to $b^n x$, $x_n := b^n x - \alpha_n$ and construct sequences $x_n^{\pm} := (\alpha_n \pm 1)b^{-n}$, which one can check both converge to x. Then,

$$\frac{f(x_m) - f(x)}{x_m - x}$$

is an infinite sum, where the first m terms are bounded in magnitude by at most

$$\pi \sum_{n=1}^{m-1} (ab)^n < \pi \frac{(ab)^m}{ab - 1}$$

using the fact that $\cos x$ has Lipschitz constant 1, and for the tail,

$$\cos(b^{n+m}\pi x_m^+) = -(-1)^{\alpha_m}, \cos(b^{n+m}\pi x_0) = (-1)^{\alpha_m}\cos(b^n \pi x_{m+1}),$$

so that

$$\sum_{n=m}^{\infty} a^n(\cos(b^n \pi x_n^+) - \cos(b^n \pi x_0) \geqslant (ab)^m \frac{1 + \cos(\pi x_{m+1})}{1 + x_{m+1}} \geqslant \frac{2}{3},$$

where we used the facts that the sum only has positive terms and took the $n = m$ term and $x_{m+1} \in (-\frac{1}{2}, \frac{1}{2}]$. These two inequalities show that

$$\frac{f(x_m^+) - f(x)}{x_m^+ - x} = (-1)^{\alpha_m}(ab)^m \eta_1 \left(\frac{2}{3} + \epsilon_1 \frac{\pi}{ab - 1}\right)$$

for $|\epsilon_1| \leqslant 1$ and $\eta_1 > 1$. An analogous argument for x_m^- shows that

$$\frac{f(x_m^-) - f(x)}{x_m^- - x} = -(-1)^{\alpha_m}(ab)^m \eta_2 \left(\frac{2}{3} + \epsilon_2 \frac{\pi}{ab - 1}\right).$$

Given the condition on ab, we have that both sides have different signs and in fact diverge to $\pm\infty$, so f is not differentiable at x. Moreover, writing

$$f_\alpha(x) = \sum_{n=1}^{\infty} b^{-n\alpha} \cos(b^n \pi x),$$

one can show that f is α-Hölder continuous for $\alpha \leqslant -\frac{\ln a}{\ln b}$. In particular, W_1 is an example of an α-Hölder continuous function for all $\alpha > 1$ that is not Lipschitz.

Example 1.4.4. Hölder continuity is a very "weak" definition of continuity. For example, the Cantor function is α-Hölder continuous with $\alpha = \frac{\log 2}{\log 3}$. As another example, take the **space-filling curves.** Let $\mathcal{C}$ be the Cantor set, considered as a topological space, and $h : \mathcal{C} \to [0,1]$ be surjective (for example, take the restriction of the Cantor function $C : \mathcal{C} \to [0,1]$). Then, since $\mathcal{C}$ is homeomorphic to $\mathcal{C} \times \mathcal{C}$, one gets a surjective map

$$\mathcal{C} \xrightarrow{\sim} \mathcal{C} \times \mathcal{C} \twoheadrightarrow [0,1] \times [0,1],$$

which may be extended to a continuous function on $[0,1]$. Note that such a map must necessarily not be injective, as it would otherwise be a homeomorphism of a unit interval and the unit square. It thus follows that space-filling curves derived from the Cantor function are Hölder continuous.

We now have many different proposed types of continuity, related by the following inclusions on $\mathbb{R}$:

$$C^1 \subset \text{Lipschitz continuous} \subset \text{AC} \subset \text{continuous and BV} \subset \text{differentiable a.e.}$$

We provide a list of relevant examples:

(a) $|x|$ is Lipschitz but not C^1 on [-1,1].

(b) $\sqrt{x}$ is AC on [0,1] since FTOC holds, but not Lipschitz, since its derivative is not bounded.

(c) Cantor's function is continuous and BV (since it is monotonic) but not AC (since its derivative is zero a.e.).

(d) $f(x) = \begin{cases} \sin(\frac{1}{x}) & x \neq 0 \\ 0 & x = 0 \end{cases}$ is differentiable a.e. on $[0,1]$ but not BV. Similarly, $x\sin(\frac{1}{x})$ is not BV (as its envelope is given by x, and the harmonic series diverges).

(e) The Weierstrass function W_1 is Hölder continuous for all $\alpha > 1$ but differentiable nowhere.

(f) $\frac{1}{\ln x}$ is AC but not Hölder continuous for any α.

Additionally, we may relate the notions of classical, weak, and distributional derivatives as follows:

(a) A function on $\mathbb{R}$ is weakly differentiable with derivative in L^1 iff it is absolutely continuous. Thus, AC is the set of functions whose derivatives are also functions. In $\mathbb{R}^n$ for $n \geqslant 2$, u is weakly differentiable iff u is **absolutely continuous on lines (ACL).** Moreover, if $u \in BV(\mathbb{R}^n)$, $u' \in L^1$.

(b) If $u \in BV(\mathbb{R}^n)$, then the distributional derivative u' is a Radon measure. If $n = 1$, then u' is classically defined a.e. and $u' \in L^1$, but the derivative is strictly weaker than the weak derivative (since it is a measure, not a function). For example, the derivative of the Cantor function is 0 a.e., while the weak derivative does not exist. However, by Radon-Nikodym, one can write $u = u_{ac} + u_j + u_s$, where $u_{ac} \in AC, u_j$ is a jump function (that is, a distribution corresponding to a discrete measure), and u_s is a singular continuous function with $u_s' = 0$ a.e. For $n \geqslant 2$, u need not even be differentiable a.e., even if u is continuous.

1.4.4 Exercises

Problem 1.4.1 (Fall 2013 Problem 12). Suppose $f : [0,1] \to \mathbb{R}$ is continuous and absolutely continuous on $(0,1]$. Show that f is not necessarily absolutely continuous on $[0,1]$, but that if it is of bounded variation on $[0,1]$, then it is absolutely continuous on $[0,1]$.

Proof. $f(x) = x \sin \frac{1}{x}$ is continuous but not BV on $[0,1]$, and absolutely continuous on $(0,1]$ since it satisfies the fundamental theorem of calculus. Now, if f is assumed to be of bounded variation, then we can consider the total variation function $T_{F(1-t)}$, which is by assumption a monotonic bounded increasing function on $[0,1]$. Thus, for any $\epsilon > 0$ there exists a $\delta > 0$ s.t. $T_{F(1-t)}(1) - T_{F(1-t)}(1-\delta) < \epsilon$, i.e. the total variation of f on $[0,\delta]$ is less than ϵ. Then, considering F on $[0,\delta]$ and $[\delta,1]$, we use absolute continuity on $[\delta,1]$ to conclude that F is absolutely continuous on $[0,1]$. ☐

Problem 1.4.2 (Fall 2016 Problem 1). Show that if $f \in L^1$ and

$$\lim_{h\to 0} \int \frac{|f(x+h) - f(x)|}{h} dx = 0,$$

then $f = 0$ a.e.

Proof. The clever trick is to use the Lebesgue Differentiation Theorem. Namely,

$$\int_c^d \frac{f(x+h) - f(x)}{h} dx = \frac{\int_c^{c+h} f(x)dx - \int_d^{d+h} f(x)dx}{h} \to 0$$

as $h \to 0$ implies that $f(c) = f(d)$ if c, d are Lebesgue points of f. But a.e. point is a Lebesgue point, so f is constant a.e., and since it is in L^1, f therefore is zero a.e. ☐

1.5 Convexity

Definition 1.5.1. A function $f : X \to \mathbb{R}$ is **convex** if $f(tx + (1-t)y) \leqslant tf(x) + (1-t)f(y)$ for all $t \in [0,1], x, y \in X$.

Theorem 1.5.1 (Geometric Hahn-Banach)**.** *If X, Y are two closed convex disjoint subsets, then there exists a hyperplane that separates X, Y.*

There is a very deep fact that relates convexity to the weak topology.

Corollary 1.5.1. *A convex set A is closed iff it is weakly closed.*

Proof. If A is weakly closed and $x_n \to x$, then $\phi(x_n) \to \phi(x)$ for all bounded functionals ϕ, i.e. $x \in A$. Conversely, if A is convex and closed, we show that A^c is weakly open. Indeed, for $a \in A^c$, by geometric Hahn-Banach there exists a separating hyperplane between a and A, which precisely implies that A^c is weakly open. ☐

Theorem 1.5.2 (Jensen's Inequality)**.** *If f is convex, then $f(\frac{1}{|X|}\int_X u(x)dx) \leqslant \frac{1}{X}\int_X f(u(x))dx$.*

Proof. Prove for sums by the definition of convexity, and pass to the limit into the integral. ☐

Definition 1.5.2. The **subdifferential** of a function $f : \mathbb{R}^n \to \mathbb{R}$ at x_0 is

$$\partial f(x_0) := \{y \in \mathbb{R} : f(x) \geqslant f(x_0) + y \cdot (x - x_0) \quad \forall x \in \mathbb{R}^n\}.$$

Proposition 1.5.1. *Given some regularity, a convex function can be characterized by the following statements:*

(a) If $f \in C^2(\mathbb{R}^n)$, f is convex iff $D^2 f$ is everywhere positive semi-definite.

(b) If $f \in C^1(\mathbb{R}^n)$, f is convex iff $f(y) \geqslant f(x) + \nabla f(x)(y - x)$, i.e. its epigraph is convex.

(c) $f : \mathbb{R}^n \to \mathbb{R}$ is convex iff for all x, $\partial f(x) \neq \varnothing$.

Proof. We prove the third statement. If f is convex, Hahn-Banach guarantees the existence of a supporting hyperplane, which defines a subdifferential. Conversely, taking $g \in \partial f(x_\alpha)$, $x_\alpha = \alpha y + (1 - \alpha)x$, since

$$f(y) \geqslant f(x_\alpha) + g \cdot (y - x_\alpha),$$
$$f(x) \geqslant f(x_\alpha) + g \cdot (x - x_\alpha),$$

multiply the first equation by α, the second by $1 - \alpha$ and add to get

$$\alpha f(y) + (1 - \alpha) f(x) \geqslant f(y_\alpha).$$

□

Lemma 1.5.1. *A convex function f attains a minimum. x is a minimum of f iff $0 \in \partial f(x)$.*

Here is an important theorem regarding the regularity of convex functions.

Proposition 1.5.2. *A convex function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable except on at most a countable set.*

Proof. For brevity we show that the set of nondifferentiability has Lebesgue measure 0. We show that convex functions are locally Lipschitz. We assume without proof that all convex functions on $\mathbb{R}^n$ are continuous. First, suppose f is bounded above on $B(x_0, \delta)$. Then, $f(x_0) = f(\frac{(2x_0 - x) + x}{2}) \leqslant \frac{1}{2} f(x) + \frac{1}{2} f(2x_0 - x)$, i.e. $f(x) \geqslant 2f(x_0) - f(2x_0 - x)$, since $2x_0 - x \in B(x_0, \delta)$, so f is bounded. Then, For $x, y \in B(x_0, \frac{\delta}{2})$, set

$$u = x + \frac{\delta}{2} \frac{(x - y)}{\|x - y\|},$$

and suppose f is bounded by M on $x, y, u \in B(x_0, \delta)$. If $\alpha = \frac{\delta}{2\|x - y\|}$,

$$x = \frac{1}{\alpha + 1} u + \frac{\alpha}{\alpha + 1} y.$$

Then,

$$f(x) - f(y) \leqslant \frac{1}{\alpha + 1} f(u) + \frac{\alpha}{\alpha + 1} f(y) - f(y) = \frac{f(u) - f(y)}{\alpha + 1} \leqslant \frac{4M}{\delta} \|x - y\|.$$

Since f is locally Lipschitz, Rademacher's theorem implies that f is differentiable a.e. □

Definition 1.5.3. A function $f : X \to \mathbb{R}$ is **lower (upper) semi-continuous** if the epigraph (hypograph) $\{(x, t) \in X \times \mathbb{R} : f(x) \geqslant (\leqslant) t\}$ is closed. Alternatively, f is lower (upper) semi-continuous at x_0 if $f(x_0) \leqslant \liminf_{x \to x_0} f(x)$ $(f(x_0) \geqslant \limsup_{x \to x_0} f(x)$.)

Remark 1.5.1. Think of upper-semicontinuous functions as functions with only discontinuities that can jump up and lower-semicontinuous functions with discontinuities that can only jump down.

Proposition 1.5.3. *(a) Indicator functions of open (closed) sets are lower (upper) semi-continuous.*

(b) Sums/products of lower/upper semi-continuous functions are lower/upper semi-continuous.

(c) Arbitrary infima (maxima) of upper (lower) semi-continuous functions are upper (lower) semicontinuous.

(d) A lower (upper) semi-continuous function on a compact set K attains a minimum (maximum).

(e) (Baire's Theorem) A lower (upper) semi-continuous function on a metric space X is the monotone limit of an increasing (decreasing) sequence of continuous functions.

Proof. We prove Baire's theorem. WLOG, suppose X is compact and f is upper semi-continuous. Define

$$f_n(x) = \sup_{y \in X}(f(y) - nd(x,y)).$$

Clearly, $f \leqslant f_n$ for all n. Note that $x \to d(x,y)$ is continuous, so f_n is the supremum of a sequence of continuous functions and therefore lower semi-continuous. Moreover, f_n is upper semi-continuous (since f is), as for $x_m \to x$,

$$\begin{aligned} f_n(x) &\geqslant \sup_y (f(y) - nd(x_m,y) + n(d(x_m,y) - d(x,y))) \\ &\geqslant \limsup_m \sup_y (f(y) - nd(x_m,y)) = \limsup_m f_n(x_m). \end{aligned}$$

Thus, f_n is continuous. Clearly, f_n is monotonic. Finally, as $n \to \infty$, it is easy to see that $f_n(x) \to x$. □

Remark 1.5.2. The main reason we care about semicontinuity is in the context of optimization problems. Consider a function $F : X \to \mathbb{R}$ on a Banach space that is bounded below. Does there exist a minimizer of this functional? Even if the functional is **coercive,** i.e. grows at ∞, we need some kind of compactness to obtain a minimizer. If X is reflexive, a minimizing sequence has a weakly convergent subsequence. Then, it suffices for F to be **weakly lower semicontinuous** for a minimizer to exist. Moreover, since F is by assumption lower semicontinuous, if F is convex, one may then conclude that F is weakly lower semicontinuous. This leads to the following lemma:

Lemma 1.5.2. *If X is a reflexive space and $F : X \to \mathbb{R}$ is a coercive, convex, lower semicontinuous functional, then there exists a minimizer for F.*

In fact, there is a partial converse to this statement.

Theorem 1.5.3 (Tonelli)**.** *If ϕ is continuous, a functional $F : u \to \int \phi(x,u)dx$ is weakly lower semi-continuous on $L^p(\mathbb{R}^n)$, $1 < p < \infty$, and weak-* lower semicontinuous on $L^\infty(\mathbb{R}^n)$ iff $u \to \phi(\cdot,u)$ is convex.*

Proof. The backward direction is immediate from the lemma above. Conversely, pick $u(x)$ to be an oscillating function between $a, b \in \mathbb{R}^m$ so that $u_n(x) := u(nx) \overset{*}{\rightharpoonup} ta + (1-t)b$, the average of u. Then, $\phi(u_n) \overset{*}{\rightharpoonup} t\phi(a) + (1-t)\phi(b)$, so on a finite measure set Ω,

$$\mu(\Omega)\phi(ta + (1-t)b) = \int \phi(ta + (1-t)b)dx \leqslant \mu(\Omega) \liminf_n \int \phi(u_n)dx = \mu(\Omega)(t\phi(a) + (1-t)\phi(b)).$$

□

1.6 Lebesgue-Radon-Nikodym Theorem

We know that $f \geqslant 0 \in L^1(X)$ defines a finite measure μ on X by $\mu(E) = \int_E f dx$, and likewise an arbitrary $f \in L^1$ defines a signed measure. Turns out, the converse of this statement is true, and a finite measure gives rise to an integrable function on a measure space.

Definition 1.6.1. Let X be a measure space, and let μ, ν be two measures. We say

$$\mu \ll \nu \quad (\mu \text{ is absolutely continuous with respect to } \nu)$$

if $\nu(E) = 0$ implies $\mu(E) = 0$. Moreover, we say

$$\mu \perp \nu \quad (\mu \text{ and } \nu \text{ are mutually singular})$$

if $X = A \cup B$, where A, B are disjoint, μ is supported on A, and ν is supported on B. A measure ν is said to be **discrete with respect to** μ if ν is supported on at most a countable set of elements, each with positive measure, and $\nu \perp \mu$. One calls a measure **singular/absolutely continuous/discrete** if it is singular/absolutely continuous with respect to the Lebesgue measure μ.

Example 1.6.1. Any discrete measure is singular. The measure given by $A \to \int_A f d\mu$, where f is the Cantor function, is an example of a non-discrete singular measure.

Definition 1.6.2. An **atom** in a measure space (X, μ) is a set A s.t. $\mu(A) > 0$, and $B \subsetneq A$ implies $\mu(B) = 0$. An atom defines an equivalence class $[A]$ where any two sets differ by a null set. If a **σ-finite** measure space consists only of atoms, it is called **atomic.**

Lemma 1.6.1. *In an atomic measure space, there are at most countably many atomic classes.*

Proof. Each atomic class is disjoint and has positive measure, so since X is σ-finite, we are done. ☐

Example 1.6.2. A measure on $[0, 1]$ that takes the value 1 on co-countable sets and 0 on countable sets is atomic but not discrete, with one atomic class.

Lemma 1.6.2 (Absolute Continuity)**.** *If μ is a **finite** signed measure and ν is a measure, then $\mu \ll \nu$ iff for all $\epsilon > 0$, there exists $\delta > 0$ such that $|\mu(E)| < \epsilon$ whenever $\nu(E) < \delta$.*

Proof. The backward direction is trivial. For the forward direction, proceed by contradiction. Then, for some $\epsilon > 0$, for all $\delta_k = \frac{1}{2^k}$, there exists an E_k such that $|\mu(E_k)| \geqslant \epsilon$ but $\nu(E_k) < \delta_k$. Then, by the Borel-Cantelli lemma, $\nu(\limsup E_k) = 0 \implies \mu(\limsup E_k) = 0$. But this is impossible since

$$\mu(\bigcap_{n=1}^{\infty} \bigcup_{k \geqslant n} E_k) \geqslant \mu(\bigcup_{k \geqslant n} E_k) \geqslant \epsilon.$$

☐

Now, notice that $f \in L^1(X)$ defines an **absolutely continuous measure** on X. Lebesgue-Radon-Nikodym states that in fact all absolutely continuous measures arise in this way.

Theorem 1.6.1 (Lebesgue-Radon-Nikodym)**.** *Let (X, ν) be a σ-finite measure space. Then, if μ is a σ-finite signed measure such that $\mu \ll \nu$, there exists $f \in L^1(X)$ such that*

$$\mu(E) = \int_E f d\nu,$$

*and such an f is defined uniquely a.e. $\frac{d\nu}{d\mu} := f$ is called the **Radon-Nikodym derivative** of μ w.r.t. ν, Moreover, if μ is not absolutely continuous, then, $\mu = \mu_1 + \mu_2$, where $\mu_1 \ll \nu, \mu_2 \perp \nu$. Additionally, $\mu_2 = \mu_d + \mu_s$, where μ_d is discrete with respect to ν and $\mu_s \perp \nu$.*

Proof. In the case of positive measures, we define

$$f = \sup\{g : \int_E g d\nu \leqslant \mu(E), E \subset X \text{ measurable}\}.$$

□

Corollary 1.6.1.
If $\nu \ll \mu \ll \rho$, *then* $\frac{d\rho}{d\nu} = \frac{d\rho}{d\mu}\frac{d\mu}{d\nu}$.

Corollary 1.6.2. *If* $\nu_1 \ll \mu_1, \nu_2 \ll \mu_2$, *then*

$$\frac{d(\nu_1 \otimes \nu_2)}{d(\mu_1 \otimes \mu_2)} = \frac{d\nu_1}{d\mu_1}\frac{d\nu_2}{d\mu_2}.$$

Example 1.6.3. For a random variable X with an absolutely continuous distribution function F,

$$\mathbb{E}[X] = \int_\Omega X dP = \int x dP^* = \int_{\mathbb{R}} x \frac{dP^*}{d\mu} d\mu = \int_{\mathbb{R}} x f(x) dx,$$

where

$$\int_a^b f(x) dx = P(X \in [a, b]).$$

is the pdf of X.

Example 1.6.4. Let $X = ([0, 1], \mu)$ and ρ be a Borel measure on X such that $\mu((a, b)) = b^2 - a^2$ and $\mu(\{0\}) = \mu(\{1\}) = 0.5$. Then,

$$\rho(E) = 0.5\chi_{1\in E} + 0.5\chi_{0\in E} + \int_E x dx,$$

so

$$\rho = \rho_1 + \rho_2, \rho_1 \ll \mu, \frac{d\rho_1}{d\mu} = x, \rho_2 = \delta_0 + \delta_1,$$

where δ_a is the Dirac delta measure at a.

1.7 Continuous Functions on a Compact Hausdorff Space

Here we list a number of important topological and measure-theoretic results that can be applied to continuous functions on a compact Hausdorff space X. Sometimes, for sake of generality, we will want to work over even more general types of sets.

Definition 1.7.1. Define a topological space X to be **locally compact Hausdorff (LCH)** if it is Hausdorff and every point has a base of compact sets.

Example 1.7.1. $\mathbb{Q}$ with the usual topology is a separable Hausdorff metric space, but is not locally compact.

Definition 1.7.2. A subset $A \subset C(X)$ **separates points** if for any $x, y \in X$ there exists $f \in A$ such that $f(x) \neq f(y)$.

Theorem 1.7.1 (Urysohn's Lemma)**.** *Let X be a compact Hausdorff space. Then, for any two disjoint closed subsets A, B of X, there exists $f \in C(X)$ such that $f(A) = 0, f(B) = 1$.*

Corollary 1.7.1. *If X is compact Hausdorff, $C(X)$ separates points.*

Definition 1.7.3. A subset $A \subset C(X)$ is **equicontinuous** if for all $\epsilon > 0$, there exists $\delta > 0$ such that for all $f \in A$, $|x - y| < \delta \implies |f(x) - f(y)| < \epsilon$.

Theorem 1.7.2 (Arzela-Ascoli)**.** *A subset of $C(X)$ is relatively compact iff it is equicontinuous and bounded.*

Corollary 1.7.2. *If $A \subset C^1(X)$ is bounded, A is relatively compact in $C(X)$.*

Theorem 1.7.3 (Stone-Weierstrass Theorem)**.** *If $A \subset C(X)$ is a unital C^*-algebra (i.e. A is closed under conjugation), A is dense in $C(X)$ iff it separates points.*

In fact, one has the following generalization:

Corollary 1.7.3 (Stone-Weierstrass)**.** *If X is a locally compact Hausdorff space, then a subalgebra $A \subset C_0(X)$ is dense iff it separates points and if there is no $x \in X$ such that A vanishes on x.*

Proof. Here is a sketch of the proof: you show that for $f, g \in A$, $|f| \in A, \min(f, g) \in A, \max(f, g) \in A$. Then, you construct a sequence of functions g_x that match f on certain points and are above it otherwise and use local compactness to cover X with those functions. Then, do the same with those functions from below to conclude. □

1.8 Riesz Represenation Theorem and Convergence of Measures

In this section, our goal is to connect the regularity properties of continuous and integrable functions, which will require some restrictions on the spaces and measures that we're dealing with. Our goal is to build up to the Riesz Representation theorem, which provides a direct description of the dual of $C(X)$.

Definition 1.8.1. A Borel measure μ is called a **Radon measure** if μ is **finite on compact subsets (f.o.c.s.)**, **inner regular on open sets,** i.e.

$$\mu(U) = \sup_{K \subset U, K \text{ compact}} \mu(K)$$

for U open, and **outer regular on Borel sets,** i.e.

$$\mu(O) = \inf_{O \subset U, U \text{ open}} \mu(U)$$

for O Borel. If μ is both inner and outer regular on Borel sets, μ is called **regular.**

Remark 1.8.1. Note that a regular measure is a slightly stronger condition than a Radon measure.

Remark 1.8.2. A Radon measure μ on a LCH space X is in fact inner regular on all σ-finite sets. Thus, a σ-finite Radon measure on an LCH space X is regular. This shows that for σ-finite measures on LCH spaces, there is no difference between Radon and regular measures.

Theorem 1.8.1. *If X is an LCH space s.t. every open set is σ-compact, then every f.o.c.s. Borel measure is Radon, and therefore also regular (since σ-compact + finite on compact subsets $\implies$ σ-finite).*

Corollary 1.8.1. *If X is a separable LCH metric space, every f.o.c.s Borel measure is regular.*

Corollary 1.8.2. *If X is a **complete** separable metric space (i.e. a **Polish space**), every f.o.c.s. Borel measure is regular.*

Remark 1.8.3. The condition that **every** open subset be σ-compact is important, as there exists a separable LCH σ-compact space for which this does not hold.

Corollary 1.8.3. *Every f.o.c.s. Borel measure on $\mathbb{R}^n$ is regular.*

Definition 1.8.2. The **(total) variation** of a signed measure μ is given by the Jordan decomposition $|\mu|(X) = \mu^+(X) - \mu^-(X)$. More generally, if μ is complex,

$$|\mu|(X) = \sup_{\pi} \sum_{A \in \pi} |\mu(A)|,$$

where π is a countable partition of X.

Definition 1.8.3. If X is a topological space, define $C_c(X) \subset C_0(X) \subset C_b(X)$ to be the spaces of compactly supported, vanishing at infinity, and bounded functions on X, all with the supremum norm. A function f is said to **vanish at** ∞ if for all $\epsilon > 0$, $|f| < \epsilon$ outside a compact set $K \subset X$.

Lemma 1.8.1. *If X is an LCH space, then $C_0(X), C_b(X)$ are Banach spaces, and the closure of $C_c(X)$ is $C_0(X)$.*

Proof. The first two claims directly follow from the fact that X is locally compact. The second is a simple consequence of Urysohn's lemma. □

Corollary 1.8.4. *If X is compact, $C(X) = C_0(X) = C_b(X) = C_c(X)$.*

Theorem 1.8.2 (Riesz-Markov-Kakutani). *If X is an LCH space, then $C_0(X)^* \cong \mathcal{M}_b(X)$, the space of complex Radon measures with finite variation on X (i.e. measures such that* $\operatorname{Re}\mu, \operatorname{Im}\mu$ *are Radon), under the equivalence*

$$\phi(f) = \int_X f(x) d\mu_\phi(x),$$

and with $\|\phi\| = |\mu_\phi|(X)$. Moreover, positive functionals on $C_0(X)$ are isometrically isomorphic to finite Radon measures.

Corollary 1.8.5. *$\mathcal{M}_b(X)$ equipped with the total variation norm is a Banach space.*

Proof. By Riesz representation, since $\mathcal{M}_b(X) = C_0(X)^*$, and a dual of a space is always Banach, we conclude. □

Corollary 1.8.6. *If every open set in X is σ-compact, then $C_0(X)^* \cong \mathcal{M}_b(X) \cong \mathcal{B}_b(X)$, the space of finite complex Borel measures on X with the total variation norm. In particular, this is true for **separable LCH metric spaces** X.*

Intuitively, we conclude that the necessity of working over a separable metric space is what makes Radon and Borel measures equivalent, and the LCH property is what is required by the Riesz Representation theorem itself.

Example 1.8.1. Any bounded linear functional on $C_0(\mathbb{R}^n)$ is given by a finite Borel measure.

Proposition 1.8.1. *If X is a compact Hausdorff metric space, $C(X)$ is separable.*

Proof. Let $A \subset X$ be a dense countable subset and consider the countable unital $\mathbb{Q}$-subalgebra generated by 1 and $h_a(x) = d(a, x)$ for $a \in A$. It clearly separates points, so by Stone-Weierstrass, it is dense in X. $\square$

Proposition 1.8.2. *If μ_n is a bounded sequence of Borel measures on a compact Hausdorff metric space X, there exists a Borel measure μ and a subsequence μ_{n_k} such that*

$$\int f d\mu_{n_k} \to \int f d\mu$$

for all $f \in C(X)$.

Proof. Since $C(X)$ is separable, by Banach-Alouglu, a bounded ball in $C(X)^* \cong \mathcal{B}(X)$ is weak-* sequentially compact. $\square$

Proposition 1.8.3. *If X is a compact metric space, the set $\mathcal{M}_1$ of probability measures on X with the weak-* topology is a compact metric space.*

Proof. $\mathcal{M}_1$ is homeomorphic to the subset of positive linear functionals of norm 1 with respect to the weak$-$* topology. Since $C(X)$ is separable, by Banach-Alaouglu, this subset is weak-* sequentially compact and the weak-* topology is in fact metrizable. Thus, $\mathcal{M}_1$ is a compact metric space. $\square$

While we have considered a lot of different kinds of convergence for functions, the equivalence of functions and measures (due to Radon-Nikodym) suggests a number of definitions for the convergence of measures.

Definition 1.8.4. One says that $\mu_n \to \mu$ **strongly or setwise** if $\mu_n(A) \to \mu(A)$ for all measurable A. One says that $\mu_n \to \mu$ **vaguely (or weak-*)** if for all $f \in C_0(X)$, $\int f d\mu_n \to \int f d\mu$. Similarly, one says that $\mu_n \to \mu$ **weakly** of for all $f \in C_b(X)$, $\int f d\mu_n \to \int f d\mu$.

Remark 1.8.4. Weak convergence is a misnomer, since $C_b(X) \neq \mathcal{M}_b^*$.

It is easy to see that strong convergence implies weak and weak-* convergence. The following theorem provides a partial converse.

Theorem 1.8.3 (Portmanteau Lemma)**.** *Given a metric space X, TFAE:*

(a) $\mu_n \rightharpoonup u$.

(b) $\liminf \mu_n(O) \geqslant \mu(O)$ *for all open $O \subset X$.*

(c) $\limsup \mu_n(K) \leqslant \mu(K)$ *for all closed $K \subset X$.*

(d) $\liminf \int f d\mu_n \geqslant \int f d\mu$ *for all lower semicontinuous bounded below f.*

(e) $\limsup \int f d\mu_n \leqslant \int f d\mu$ *for all upper semicontinuous bounded above f.*

(f) $\lim \mu_n(A) = \mu(A)$ *for all A with $\mu(\partial A) = 0$.*

Proof. we show $(a) \implies (c)$. For K closed, define $K_n = \{x : d(K, x) \leqslant \frac{1}{n}\}$, and let F_k be a continuous function that is 1 on K and 0 on K_n^c. Then,

$$\limsup_n \mu_n(K) \leqslant \limsup_n \int F_k d\mu_n \to \int F_k d\mu \leqslant \mu(K_n).$$

Taking $n \to \infty$ completes the proof. By taking complements, one sees that (b) is equivalent to (c). Using Baire's theorem to monotonically approximate upper (lower) semiicontinuous functions yields equvialence with (d) and (e). Together, (b) and (c) imply (d), since

$$\limsup \mu_n(\overline{A}) \leqslant \mu(\overline{A}), \liminf \mu_n(A^o) \geqslant \mu(A^o).$$

To see that $(b) \implies (a)$, note that by Fatou,

$$\liminf \int f d\mu_n = \liminf \int_0^\infty \mu_n\{x > \lambda\} \geqslant \int_0^1 \liminf \mu_n\{f > \lambda\} \geqslant \int \mu\{f > \lambda\} = \int f d\mu$$

and replacing f with $-f$ completes the proof. □

Definition 1.8.5. For $f : (X, \mu) \to (Y, \nu)$, the **pushforward measure associated to** f on Y is $\mu_f(B) := \mu(f^{-1}(B))$.

Example 1.8.2. If $f(x) = c$ is constant, μ_f is the Dirac measure on c.

Remark 1.8.5. The defining property of pushforward measures for $g : Y \to Z$ is

$$\int g \circ f \mu = \int g d\mu_f$$

Definition 1.8.6. We say $f_n \to f$ **in distribution** if $\mu_{f_n} \rightharpoonup \mu_f$.

In fact, we have an analogue of Arzela-Ascoli for measures.

Theorem 1.8.4 (Prokhorov's Theorem)**.** *Let S be a separable metric space, and $\mathcal{M}_1(S)$ be the space of Borel probability measures on S. Then, a subset $A \subset \mathcal{M}_1(S)$ is weakly precompact iff it is tight. Moreover, if S is complete, then the weak topology is completely metrizable.*

Proof. If A is tight, the rougly speaking, one can pick a countable subsequence of sets and use a diagonal argument to find a convergent subsequence on those sets, and extend it to a weak limit by regularity of the measure.

Conversely, suppose A is weakly precompact. Note that $K \subset S$ is compact iff

$$K = \bigcap_j \bigcup_{i=1}^{N_j} B(x_i, \frac{1}{j}),$$

where $\{x_i\}$ is a countable dense subset of S. If for every j we can find an N_j such that

$$\mu\left(\bigcup_{i=1}^{N_j} B(x_i, \frac{1}{j})\right) > 1 - (1 - \frac{1}{2^j})\epsilon$$

for all $\mu \in A$, then K as above would satisfy $\mu(K^c) < \epsilon$ for all $\mu \in A$. If not, then there exists a j such that for all N_j there is a sequence of measures $\mu_k \overset{*}{\rightharpoonup} \nu$ such that

$$\mu_k\left(\bigcup_{i=1}^{N_j} B(x_i, \frac{1}{j})\right) \leqslant 1 - (1 - \frac{1}{2^j})\epsilon.$$

But picking N_j to cover S in the limit, we then get $\nu(S) \leqslant 1 - (1 - \frac{1}{2^j})\epsilon$, a contradiction. □

1.9 Measure Theory

Lemma 1.9.1 (Borel-Cantelli). *(a) If $\sum \mu(A_n) < \infty$, then $\mu(\limsup A_n) = 0$.*

(b) If μ is a probability measure, $\sum \mu(A_n) = \infty$, and $\mu(A_k \cap A_j) = \mu(A_k)\mu(A_j)$, then $\mu(\limsup A_n) = \mu(X)$.

Proof. We prove (b). Indeed, it suffices to show that $\limsup_n \mu(\bigcap_{n \geqslant N} A_n^c) = 0$, as

$$\mu(\limsup A_n) = \mu\left(\bigcap_N \bigcup_{n \geqslant N} A_n\right) \geqslant \liminf \mu\left(\bigcup_{n \geqslant N} A_n\right) = 1 - \limsup_n \mu\left(\bigcap_{n \geqslant N} A_n^c\right).$$

Now, since the A_n are indepenent, one can easily check that

$$\mu(A_k^c \cap A_j^c) = \mu(X) - \mu(A_k \cup A_j) = \mu(X) - \mu(A_k) - \mu(A_j) + \mu(A_k)\mu(A_j) = \mu(A_k^c)\mu(A_j^c),$$

so $\mu(\bigcap_{n \geqslant N} A_n^c) = \prod_{n=N}^{\infty}(1 - \mu(A_n)) \to 0$, since $\sum \log(1 - \mu(A_n)) \sim -\sum \mu(A_n) = -\infty$. ☐

Here are some common counterexamples used in measure theory:

Example 1.9.1. (a) The Cantor set $\mathcal{C}$ - it is a closed nowhere dense subset of $[0, 1]$ of measure zero.

(b) The fat Cantor set $\mathcal{C}_\alpha$ - it is a closed nowhere dense subset of $[0, 1]$ of measure $\alpha \in (0, 1)$.

(c) Define the **Cantor function** as follows - let $c : \mathcal{C} \to [0, 1]$ be defined by replacing all the 2's in the expansion of a number with 1's and extending the function to be locally constant on the remaining intervals. Then, f is monotonic uniformly continuous (in fact, Hölder continuous) but not absolutely continuous.

(d) If $C(x) = c(x) + x$, then $C(x)$ is a homeomorphism between $[0, 1]$ and $[0, 2]$, as it is a bijective continuous map from a compact to a Hausdorff space. In particular, f maps Borel sets to Borel sets, and if $N \subset C(\mathcal{C})$ is a nonmeasurable set, $f^{-1}(N) \subset \mathcal{C}$ is a Lebesgue measurable set, but not Borel measurable set, as $f(f^{-1}(N)) = N$ is not Borel. Moreover, $\chi_{f^{-1}(N)}$ is a Lebesgue but not Borel measurable function.

(e) Every subset of a null set is Lebesgue measurable since the Lebesgue measure is complete. Moreover, by an analogue of the Vitali set construction, every positive measure set contains a nonmeasurable subset.

How does one formally construct a measure? That is the question answered by the Caratheodory theorem.

Theorem 1.9.1 (Caratheodory Extension Theorem). *Let $\mathcal{A}$ be an **algebra** (i.e. closed under finite intersections and complements) of subsets of a set X. Then, a **premeasure** (that is, a measure on the algebra) ν extends to an **outer measure** μ^* on $\mathcal{P}(X)$, which restricts to a measure μ on the σ-algebra of μ-measurable sets, i.e. sets A for which*

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

for all $E \in \mathcal{P}(X)$. Moreover, if ν is σ-finite, μ is unique.

Example 1.9.2. If μ is the counting measure on $\mathbb{R}$ and ν is the infinite measure, then they agree on all cofinite sets, but not on the Borel σ-algebra.

If two measures agree on a set that generates a σ-algebra, do they agree on that σ-algebra? Turns out the answer is yes under certain conditions. There are two theorems that allow us to relax our hypotheses.

Theorem 1.9.2 (Monotone Class Theorem). *If $\mathcal{A}$ is an algebra closed under countable increasing unions and countable decreasing intersections, then it is a σ-algebra.*

Theorem 1.9.3 (Dynkin's $\pi - \lambda$ theorem). *If $\mathcal{A}$ is a π-system, i.e. a class closed under finite intersections, then the σ-algebra it generates is the same as its* ***Dynkin class,*** *i.e the class generated by disjoint unions and complements in $\mathcal{A}$.*

Proposition 1.9.1. *If two finite measures μ, ν coincide on a class $\mathcal{C}$ closed with respect to finite intersections, then they coincide on the σ-algebra $\mathcal{B}$ it generates.*

Proof. The measures agree on a class generated by $\mathcal{C}$ generated by disjoint unions and complements. □

Corollary 1.9.1. *Two* ***finite*** *measures on a topological space that agree on all open (closed) sets agree everywhere. In particular, the Lebesgue measure on $\mathbb{R}^n$ is unique.*

Remark 1.9.1. One **cannot relax the assumption to σ-finite measures.** As a counterexample, consider

$$m(A) = |A \cap \mathbb{Q}|, n(A) = |A \cap (\mathbb{Q} \cup \{\sqrt{2}\})|$$

with respect to the counting measure $|\cdot|$. Since $\mathbb{Q}$ is countable, these measures are σ-finite and they agree on the algebra of half-open intervals, but clearly not on all Borel sets, since the restriction to the half-open intervals is not σ-finite. This is because

Theorem 1.9.4 (Disintegration Theorem). *Let X, Y be two* ***Radon spaces*** *(i.e. spaces where every finite Borel measure is Radon), $\mu \in \mathcal{M}_1(Y)$, $\pi : Y \to X$ be a measurable function, and $\nu = \mu \circ \pi^{-1} \in \mathcal{M}_1(X)$ be the pushforward measure. Then, there exists a family of probability measures $\mu_x \in \mathcal{M}_1(Y)$ for $x \in X$ such that μ_x is supported on $\pi^{-1}(x)$ and*

$$\int_Y f(y) d\mu(y) = \int_X \int_{\pi^{-1}(x)} f(y) d\mu_x(y) d\nu(x).$$

Roughly, Y should be thought of as being "parametrized by X," with π being the projection map.

Proof. Note that for $B \subset Y, A \subset X$ measurable, the disintegration formula should satisfy

$$\int_{\pi^{-1}(A)} \chi_B d\mu = \int_A \mu_x(B) d\nu.$$

Using the Lebesgue differentiation theorem, we can then extract μ_x by defining

$$\mu_x(B) := \lim_{\epsilon \to 0} \frac{1}{\nu(A_\epsilon)} \int_{\pi^{-1}(A_\epsilon)} \chi_B d\mu = \lim_{\epsilon \to 0} \frac{1}{\mu(\pi^{-1}(A_\epsilon))} \int_{\pi^{-1}(A_\epsilon)} \chi_B d\mu$$

over neighborhoods A_ϵ that shrink to x. □

Often times, we are interested in measuring the "dimension" of a set. We define the dimension according to how scaling the object affects its measure. This motivates the following definition.

Definition 1.9.1. Let X be a metric space. Define

$$H^d_\delta(S) = \inf \sum_{i=1}^{\infty} (\operatorname{diam} U_i)^d$$

over all countable covers of S by sets with $\operatorname{diam} U_i < \delta$. We define the **d-dimensional Hausdorff measure** to be the Borel measure obtained from the Caratheodory restriction of the outer measure $H^d(S) = \sup_{\delta > 0} H^d_\delta(S)$.

Remark 1.9.2. Restricting to certain classes of sets U_i (like open or closed) might change the measures but does not change the dimension of a set.

Remark 1.9.3. For $d \in \mathbb{N}$ one has $\lambda^d = \beta_d H^d$, where λ^d is the d-dimensional Lebesgue measure and β_d is the volume of the d-dimensional unit ball.

Definition 1.9.2. For every set $S \subset X$, there exists a unique $d \in [0, \infty]$ s.t. $H^{d'}(S) = 0$ for $d' > d$ and ∞ for $d' < d$. We call d' the **Hausdorff dimension** of S.

1.10 Integral Transforms

Of particular interest are various objects defined in terms of integrals. In this section, we present a few key example and how one may establish properties of these objects using real and complex analysis techniques.

Definition 1.10.1. A **integral transform** is a map of the form

$$T(f)(x) = \int K(x, y) f(y) dy,$$

where $K(x, y)$ is called the **kernel** of the transform.

Typically, one wants to establish that a certain integral transform is well-defined, and then that it is a bounded map between, say, L^p spaces. To establish these kinds of general facts, one needs the following lemmas.

Lemma 1.10.1 (Minkowski's Inequality)**.**

$$\left\| \int f(\cdot, y) dy \right\|_{L^p(dx)} \leqslant \int \|f(\cdot, y)\|_{L^p(dx)} dy.$$

Proof. By Hölder,

$$\int \left[\int f(x, y) dy \right] |g(x)| dx = \int \int f(x, y) |g(x)| dx dy \leqslant \int \|f(\cdot, y)\|_{L^p(dx)} \|g\|_q dy,$$

so the inequality follows from the fact that $\|h\|_p = \sup \int hk dx$ for $k \in L^q$, $\|k\|_q = 1$. $\square$

Example 1.10.1. (a) The Fourier/Laplace transforms are integral transforms with kernels $e^{-2\pi i \xi \cdot x}$, e^{-ts}.

(b) The Poisson integral formula is given by an integral transform with the Poisson kernel $P_r(\theta) = \frac{1-r}{1-2r\cos\theta + r^2}$.

(c) The convolution $f * g$ can be thought of as an integral transform with kernel f or g, respectively.

1.10.1 The Gamma Function

Consider the **gamma function** defined by

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt.$$

First note that this integral converges for $\operatorname{Re} z > 0$, since

$$z\int_0^\infty t^{z-1}e^{-t}dt = \int_0^\infty t^z e^{-t}dt = \int_0^1 t^z e^{-t}dt + \int_1^\infty t^z e^{-t}dt,$$

where the second term is bounded by $\int_1^\infty t^n e^{-t}dt < \infty$ for $n \in \mathbb{N}$, and the first term is clearly bounded as well. Moreover, one immediately sees that $\Gamma(1) = 1$ and $z\Gamma(z) = \Gamma(z+1)$. We thus note that $\Gamma(z) = (z-1)!$. Now, interestingly, one may attempt to define $\Gamma(z)$ for $\operatorname{Re} z \leqslant 0$ by the above relation $\Gamma(z-1) = \frac{\Gamma(z)}{z-1}$. Then, $\Gamma(z)$ is well-defined everywhere except at the nonpositive integers $-n, n \geqslant 0$.

Lemma 1.10.2. *$\Gamma : \mathbb{C} \to \mathbb{C}$ is meromorphic with simple poles at $-n, n \geqslant 0$.*

Proof. Since we are asked to prove holomorphicity, we use Morera's theorem. Clearly, Γ is continuous everywhere where it is defined. Moreover, Γ is holomorphic in the right half-plane by Fubini and Cauchy's theorem, since

$$\int_\Delta \int_0^\infty t^{z-1}e^{-t}dtdz = \int_0^\infty e^{-t}\int_\Delta e^{(z-1)\ln t}dzdt = 0.$$

Then, we proceed strip by strip, as it inductively suffices to show Γ is meromorphic in $\operatorname{Re} z > -1$. Any triangle Δ in $\operatorname{Re} z > -1$ not containing 0 may be split into triangles contained in $-1 < \operatorname{Re} z < 0$ and triangles contained in $\operatorname{Re} z > 0$ with one of the sides within ϵ of the boundary $\operatorname{Re} z = 0$. The contour integrals match on the boundary by continuity, and by Cauchy's theorem the integral in the strip and the right-half plane vanish, so by Morera's theorem, we are done. To show the singularities at $-n$ are simple poles, notice that

$$\lim_{z\to -n^-}(z+n)\Gamma(z) = \lim_{z\to -n^+}(z+n)\Gamma(z) = \lim_{z\to -n}\frac{\Gamma(z+n+1)}{(z+n-1)...(z+1)z} = \frac{(-1)^n}{n!},$$

showing that the poles at $-n$ are simple. ☐

Proposition 1.10.1. $\Gamma(z) = 1 - \frac{1}{z} + \sum_{n=1}^\infty (-1)^n\left(\frac{1}{n!(z+n)} - \frac{1}{n\cdot n!}\right)$

Proof. Mittag-Leffler. ☐

Proposition 1.10.2 (Reflection Principle)**.** $\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z}$ *for* $z \notin \mathbb{Z}$.

Proof. By Fubini, if $u = s+t, v = \frac{t}{s}$,

$$\begin{aligned}\Gamma(z)\Gamma(1-z) &= \int_0^\infty t^{z-1}e^{-t}dt\int_0^\infty s^{-z}e^{-s}ds = \int_0^\infty\int_0^\infty \left(\frac{t}{s}\right)^z e^{-(s+t)}t^{-1}dtds \\ &= \int_0^\infty\int_0^\infty \frac{v^{z-1}}{1+v}e^{-u}dudv = \frac{\pi}{\sin \pi z},\end{aligned}$$

where the latter integral may be evaluated by contour methods. Namely, if $0 < \operatorname{Re} v < 1$, the integral converges absolutely, so we integrate along a keyhole contour with a branch cut on the positive real axis. At -1, the residue is $(-1)^{z-1} = e^{\pi i(z-1)} = -e^{\pi iz}$. On the large and small circles of radii $R, \epsilon > 0$, the function is asymptotically like $\frac{R^{z-1}}{1+R} \to 0$ as $R \to \infty$ and $\frac{\epsilon^{z-1}}{1+\epsilon} \to 0$ as $\epsilon \to 0$. Finally, on the other side of the keyhole, the integral is

$$\int_0^\infty \frac{e^{(z-1)(\log|v|+(i(\theta+2\pi))}}{1+v} dv = -e^{2\pi iz} I,$$

where I is the value of the desired integral. Thus,

$$I(1-e^{2\pi iz}) = -2\pi i e^{-\pi iz} \implies I = \frac{2\pi i}{e^{\pi iz} - e^{-\pi iz}} = \frac{\pi}{\sin \pi z}.$$

□

Corollary 1.10.1. $\Gamma(z)$ *has no zeros, i.e.* $\frac{1}{\Gamma(z)}$ *is entire.*

Corollary 1.10.2. $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Proposition 1.10.3 (Stirling's Formula)**.** $\Gamma(n) \sim \sqrt{2\pi n}(\frac{n}{e})^n$.

1.10.2 Hilbert Transform

Definition 1.10.2. The **Cauchy principal value (p.v.)** of a function f with a singularity at a point b or at ∞ is defined as

$$p.v. \int_{-\infty}^{\infty} f(x)dx = \lim_{\epsilon \to 0} \lim_{\eta \to \infty} \int_{b-\eta}^{b-\epsilon} f(x)dx + \int_{b+\epsilon}^{b+\eta} f(x)dx.$$

Note that $p.v. \int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{\infty} f(x)dx$ if f is Lebesgue integrable.

Definition 1.10.3. A function f with a well-defined Cauchy principal value over all smooth compactly supported test functions defines a distribution $p.v.(f) : C_c^\infty \to \mathbb{R}$ given by

$$p.v.(f)(\phi) = \text{p.v.} \int_{\mathbb{R}} f(x)\phi(x)dx.$$

Example 1.10.2. $p.v.(\frac{1}{x}) \in \mathcal{S}'$, for

$$\int_{-\epsilon}^{\epsilon} \frac{\phi(x)}{x} dx \leqslant 2\epsilon \|\phi'\|_\infty$$

and

$$\int_{|x| \geqslant 1} \frac{\phi(x)}{x} dx \ll C\|x\phi\|_\infty.$$

Proposition 1.10.4. *For* $f, \phi \in L^1 \cap C(\mathbb{R})$*, even,* $\phi(0) = 1$*, we have*

$$p.v. \left(\frac{1}{x}\right) f = \int \frac{f(x) - f(0)\phi(x)}{x} dx.$$

Proof. Follows directly from definition. □

Theorem 1.10.1 (Sokhotski-Plemelj Formulae)**.** *Let ϕ be a Hölder continuous function defined on a closed curve $C \subset \mathbb{C}$. Then, the Cauchy integral*

$$f(z_0) = \int_C \frac{\phi(z)}{z - z_0} dz$$

defines a holomorphic function $f \in H(\mathbb{C}\backslash C)$, with limits $f_\pm$ as $z \to C$ from inside/outside equal to

$$f_\pm(z) = \pm\frac{1}{2}\phi(z) + \frac{1}{2\pi i} p.v. \int_C \frac{f(t)}{t - z} dt.$$

In particular, on the real line one has

$$\lim_{\epsilon \to 0} \int \frac{f(x)}{x \pm i\epsilon} dx = \mp i\pi f(0) + p.v. \int \frac{f(x)}{x}.$$

Proof. Defining $\phi \equiv 2\pi i f(0)$ yields the second claim from the first one. ☐

Definition 1.10.4. The **Hilbert transform** of u is defined as

$$H(u)(x) = \frac{1}{\pi} p.v. \int_{-\infty}^{\infty} \frac{u(y)}{x - y} dy := u * \frac{1}{\pi x},$$

where the convolution is taken in the sense of tempered distributions.

Proposition 1.10.5. *$\widehat{H}$ is Fourier multiplier with symbol $-i\chi_{(0,\infty)}$.*

Proof. ☐

It turns out that the Hilbert transform is a bounded operator on L^p, but the proof technique can be generalized to a much larger class of singular integral operators of convolution type.

Definition 1.10.5. An integral operator of convolution type with kernel $K \in L^1_{loc}(\mathbb{R}^n)$ is said to be of **Calderon-Zygmund type** if

$$\widehat{K} \in L^\infty, K \in C^1(\mathbb{R}^n \backslash \{0\}), |\nabla K(x)| \ll |x|^{-(n+1)}.$$

Theorem 1.10.2. *A Calderon-Zygmund type operator is of strong type (p, p) and weak type $(1, 1)$ for $1 < p < \infty$.*

2 Harmonic Analysis

2.1 Interpolation

We start with a review of some key results regarding interpolation in L^p spaces. The overarching idea is that if a function belongs to two different L^p spaces, it actually belongs to all those in between the two.

Proposition 2.1.1. *For all $1 \leqslant p \leqslant q \leqslant r \leqslant \infty$,*

$$L^p \cap L^r \hookrightarrow L^q \hookrightarrow L^p + L^r$$

are continuous inclusions, where $L^p \cap L^r$ is a Banach space with norm $\|f\|_p + \|f\|_r$ and $L^p + L^r$ is a Banach space with norm $\inf_{g+h=f \in L^q} \|g\|_p + \|h\|_r$.

Proof. That all of these are Banach spaces is left as an exercise. Note that for some $\lambda \in [0,1]$,

$$q = \lambda p + (1-\lambda)r,$$

so by Young's inequality,

$$\int |f|^q dx = \int |f|^{\lambda p+(1-\lambda)r} dx \leqslant \|f^{\lambda p}\|_{\frac{1}{\lambda}} \|f^{(1-\lambda)r}\|_{\frac{1}{1-\lambda}} \leqslant \|f\|_p^{\lambda} \|f\|_r^{1-\lambda} \leqslant \lambda \|f\|_p + (1-\lambda)\|f\|_r.$$

Similarly, if $f \in L^q$, f

$$\inf_{g+h=f\in L^q} \|g\|_p + \|h\|_r \leqslant \|\chi_{|f|>1} f\|_p + \|\chi_{|f|\leqslant 1} f\|_r \leqslant \|f\|_q.$$

□

The conclusion is that one may interpolate between intermediate L^p spaces. We now considerably generalize this approach by introducing interpolation between bounded operators.

Definition 2.1.1. f is in **weak** L^p if

$$\mu\{x : |f(x)| > \lambda\} \leqslant \frac{C^p}{\lambda^p},$$

where the smallest such C is the weak L^p norm $\|f\|_{p,w}$. We say a bounded operator $T : L^p \to L^q$ is of **strong type** (p,q), and a bounded operator $T : L^p \to L^{q,w}$ is of **weak type** (p,q).

Remark 2.1.1. Clearly, if $L^p \hookrightarrow L^{p,w}$ is continuous with norm 1.

Theorem 2.1.1 (Riesz-Thorin Interpolation Theorem)**.** *Let $T : L^{p_0} \to L^{q_0}, L^{p_1} \to L^{q_1}$ be a linear bounded operator. Then, it is also bounded as an operator $T : L^{p_t} \to L^{q_t}$, where $0 < t < 1$ and*

$$\|T\|_{q_t} \leqslant \|T\|_{p_0}^{1-t} \|T\|_{p_1}^{t}, \quad \frac{1}{p_t} = \frac{1-t}{p_0} + \frac{t}{p_1}, \quad \frac{1}{q_t} = \frac{1-t}{q_0} + \frac{t}{q_1}.$$

Remark 2.1.2. We call the **Riesz diagram** of an operator T to be the set of points (p,q) in the unit square such that T is of type $(\frac{1}{p}, \frac{1}{q})$, and the theorem tells us that such a set is convex.

Proof. The proof requires the use of the following lemma:

Lemma 2.1.1 (Hadamard Three-Lines Lemma)**.** *Let $S = \{z : 0 \leqslant Re\ z \leqslant 1\}$ and suppose $F : S \to \mathbb{C}$ is bounded, analytic on the inside, and continuous on the boundary. Then, if $M_\theta = \sup_{Re\ z=\theta} |F(z)|$, one has $M_\theta \leqslant M_0^{1-\theta} M_1^{\theta}$.*

Proof. Without loss of generality, suppose $M_0 = M_1 = 1$ (otherwise, divide by appropriate powers). Then, note that $F_n(z) = F(z)e^{\frac{z^2-1}{n}}$ converges normally to F and is bounded by 1 on the boundary, so by maximum modulus, F is bounded by 1. Note that F_n are needed to converge to 0 since one can only apply maximum modulus on a bounded set. □

Corollary 2.1.1. *One obtains the same result if one relaxes boundedness to boundedness on the boundary and instead uses the estimate $f(z) = e^{e^{O(|Im z|)}}$ on the inside.*

Corollary 2.1.2. *By setting $g(z) = F(e^z)$, one obtains the **Hadamard three-circles lemma**, which states that on an annulus, if $M(s) = \sup_{|z|=e^s} |g(z)|$, then $\log M(s)$ is convex, i.e. $\log M(r)$ is convex as a function of $\log r$.*

We proceed with the proof of Riesz-Thorin. We may normalize all operator norms to be 1. By Hölder, we obtain that

$$\int |(Tf)g| \leqslant \|f\|_{p_0}\|g\|_{q_0'}, \quad \int |(Tf)g| \leqslant \|f\|_{p_1}\|g\|_{q_1'}.$$

We claim that this holds for all $\|f\|_{p_t}\|g\|_{q_t}$ on the right. Define

$$F(s) = \int T(\mathrm{sign}(f)|f|^{p_s'})\mathrm{sign}(g)|g|^{q_s'},$$

where p_s', q_s' linearly interpolate in s between $\frac{p_t}{p_0}(\frac{q_t}{q_0})$ and $\frac{p_t}{p_1}(\frac{q_t}{q_1})$, respectively. Clearly, $F(s)$ is holomorphic, satisfies the exponential estimate, and $F(0), F(1)$ are bounded on vertical lines in the complex plane. Thus, by the Hadamard Three-Lines lemma, one obtains the claim, with the estimate holding on functions of finite support. By standard density arguments one may pass to all functions and use duality to conclude the operator norm estimates. □

Finally, we need a more powerful extension of the Riesz-Thorin interpolation theorem to weak L^p spaces.

Definition 2.1.2. A **sublinear operator** T satisfies $|T(f+g)| \leqslant |Tf| + |Tg|$ and $|T(f-g)| \geqslant |Tf - Tg|$.

Remark 2.1.3. Maximal operators, such as the Hardy-Littlewood maximal operator H_f, are suprema of linear operators and thus sublinear.

Theorem 2.1.2 (Marcinkiewicz Interpolation Theorem)**.** *If T is a sublinear operator is of weak type (p_0, q_0) and (p_1, q_1), then T is of strong type (p_θ, q_θ) for*

$$\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \frac{1}{q_\theta} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}.$$

Proof. By density arguments, its enough to work with simple functions. Let f be a simple function and let $f = \sum f_m := \sum_m f\chi_{2^m \leqslant |f| \leqslant 2^{m+1}}$ be a dyadic decomposition of f. By the layer-cake decomposition, it suffices to show that

$$\int_0^\infty \mu(|Tf| > \lambda)\lambda^{q_\theta - 1} \lesssim \|f\|_{p_\theta}^{q_\theta},$$

and by dyadic decomposition and normalizing the norm, this is equivalent to

$$\sum_{n\in\mathbb{Z}} \mu(|Tf| > 2^n)2^{q_\theta n} \lesssim 1.$$

Note that $\|f\|_{p_\theta} = 1$ is equivalent to $\sum_{m\in\mathbb{Z}} \mu(|Tf| > 2^m)2^{q_\theta m} \lesssim 1$. From sublinearity we have

$$\mu(|Tf| \geqslant 2^n) \leqslant \sum_m \mu(|f_m| \geqslant c_{nm}2^n), \quad \sum_m c_{nm} = 1.$$

By applying our two hypotheses to $\mu(|f_m| \geqslant c_{nm}2^n)$ and using the definition of each f_m one may obtain the bound

$$\mu(|Tf_m| > c_{nm}2^n) \lesssim c_{nm}^{-q_i}2^{-nq_i}2^{mq_i}\mu(|f| > 2^m)^{\frac{q_i}{p_i}}, \quad i = 0, 1.$$

It thus suffices to show that

$$\sum_n 2^{nq_\theta} \sum \min_{i=0,1} c_{nm}^{-q_i} 2^{-nq_i} 2^{mq_i} \mu(|f| > 2^m)^{\frac{q_i}{p_i}} \lesssim 1.$$

Since $p_i \leqslant q_i$ and $\sum_m \mu(|f| > 2^m) \lesssim 1$, we only need to find c_{nm} such that

$$\sum_n \min_{i=0,1} c_{nm}^{-q_i} 2^{(n\alpha q_\theta - mp_\theta)q_i x_i} \lesssim 1$$

for $\frac{1}{p_i} - \frac{1}{p_\theta} = x_i, \frac{1}{q_i} - \frac{1}{q_\theta} = \alpha x_i$. But then, if c_{nm} is a sufficiently small multiple of the above power of 2, the above sum is a geometric series, and we are done. □

Theorem 2.1.3 (Schur's Test)**.** *If T is an integral kernel operator with $\|K(x, \cdot)\|_{q_0} \leqslant B_0$ and $\|K(\cdot, y)\|_{p_1'} \leqslant B_1$. Then, T is of strong type (p_t, q_t) with norm at most $B_0^t B_1^{1-t}$ for $0 < t < 1$ with $p_0 = 1, q_1 = \infty$.*

Proof. The hypotheses and Minkowski show that T is of strong type (p_0, q_0) and (p_1, q_1), so the claim follows immediately from Riesz-Thorin. □

Remark 2.1.4. An application of Marcinkiewicz yields an analogous version of weak Schur's test.

Proposition 2.1.2 (Hardy-Littlewood-Sobolev)**.** *If $I_a f := f * |x|^{-\alpha}$, then $I_\alpha : L^p \to L^r$ for $\frac{1}{p} + \frac{\alpha}{n} = \frac{1}{r} + 1$.*

Proof. $|x|^{-\alpha} \in L^{\frac{n}{\alpha}, w}$, so one concludes by the weak version of Young's convolution inequality. □

Remark 2.1.5. Note that $\widehat{I_\alpha f} = \partial^\alpha \widehat{f}$, so this yields a well-defined fractional differentiation operator on the Fourier side.

2.2 Fourier Transform

Here are some key results about the Fourier transform that are tested quite frequently on the qual:

Definition 2.2.1. The **Fourier transform** $\widehat{f}$ of $f(x)$ is defined as

$$\mathcal{F}\{f\}(\xi) = \widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i \xi \cdot x} dx.$$

The **inverse Fourier transform** $\check{f}$ of $f(\xi)$ is defined as $\mathcal{F}^{-1}\{f\}(\xi) = \check{f}(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} f(\xi) e^{2\pi i \xi \cdot x} dx$.

(a) $\widehat{f}(0) = \int f(x)$.

(b) $\widehat{f'}(\xi) = -2\pi i \xi \widehat{f}(\xi)$.

(c) $\widehat{f * g} = \widehat{f}\widehat{g}$.

(d) $f(\widehat{x + x_0}) = e^{2\pi i x_0 \cdot \xi} \widehat{f}(\xi)$.

(e) $\mathcal{F}^2\{f\}(x) = f(-x)$, i.e. $\mathcal{F}^4 = I$, so $\mathcal{F}$ has eigenvalues $i, -1, -i, 1$.

(f) $\mathcal{F}\{e^{-\pi\|x\|^2}\} = e^{-\pi\|\xi\|^2}$.

(g) **Riemann-Lebesgue Lemma:** The Fourier transform is a linear operator $\mathcal{F} : L^1 \to C_0$, where C_0 is the space of (uniformly) continuous functions vanishing at infinity.

(h) **Plancherel's Theorem:** $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ is an isomorphism and $\mathcal{F} : L^2 \to L^2$ is a unitary isometric isomorphism. In particular, $\|f\|_2 = \|\widehat{f}\|_2$ and $\langle \widehat{f}, \widehat{g} \rangle = \langle f, g \rangle$.

Remark 2.2.1. One can also take the (inverse) Fourier transform on the torus $\mathbb{T}^k$, which is equivalent to taking the transform of a periodic function. This is known as a **Fourier series.** In this case,

$$\widehat{f}(n) = \int_{\mathbb{T}^n} f(x)e^{-2\pi i n\cdot x}dx, f(x) = \sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{2\pi i n\cdot x}$$

for $n \in \mathbb{Z}^k$. Then, the same theorems apply, except that now, $\mathcal{F} : L^2 \to l^2$ and $\mathcal{F} : L^1 \to c_0$.

Remark 2.2.2. While $\mathcal{F} : L^2 \to L^2$ is an isomorphism, it is not surjective as a map $\mathcal{F} : L^1 \to C_0$. For if it were, it would induce an isomoprhism of the dual space $\mathcal{F}^* : \mathcal{M}_b \to L^\infty$, which is not surjective since the Fourier transform of a measure is necessarily uniformly continuous. However, by a standard application of Stone-Weierstrass, the range is dense in C_0.

Remark 2.2.3. Note that the Fourier transform interchanges derivatives and multiplication. Consequently, regularity on one side implies decay on the other side and vice-versa.

Corollary 2.2.1 (Hausdorff-Young Inequality). *By Riesz-Thorin, since $\mathcal{F} : L^1 \to L^\infty$ and $\mathcal{F} : L^2 \to L^2$ is bounded, it is also bounded as an operator $\mathcal{F} : L^p \to L^q$, where $1 \leqslant p \leqslant 2, 2 \leqslant q \leqslant \infty$, and*

$$\frac{1}{p} = 1 - \frac{t}{2}, \frac{1}{q} = \frac{t}{2} \quad \Longrightarrow \quad \frac{1}{p} + \frac{1}{q} = 1,$$

i.e. $\mathcal{F} : L^p \to L^{p'}$, for $1 < p < 2$ and p, p' conjugates. In particular, applying to Fourier series yields $\mathcal{F} : L^p \to l^{p'}$ and $\mathcal{F}^{-1} : l^p \to L^{p'}$ for $1 \leqslant p \leqslant 2$.

Of particular interest are the Fourier transforms of certain compactly supported functions, which can be holomorphically extended to the upper half-plane and are summarized in the following theorems.

Theorem 2.2.1 (Paley-Wiener I). *$f \in L^2((0, \infty))$ iff $\widehat{f}$ is holomorphic in the upper half-plane and the L^2 norm of $\widehat{f}$ is uniformly bounded over horizontal lines.*

Theorem 2.2.2 (Payley-Wiener II). *$f \in L^2(\mathbb{R})$ is compactly supported in $[-A, A]$ iff $\widehat{f}$ is holomorphic in the upper half-plane and of exponential type A.*

Proof. If $f \in L^2(\mathbb{R})$ has compact support in $[-A, A]$, for all $\widehat{f}(\xi)$ is well-defined in the upper half plane (as one has a decaying exponential). Moreover, by Fubini and Cauchy's theorems, one may check that $\widehat{f}$ is holomorphic. Finally,

$$f(a + bi) = \int_{-A}^{A} f(x)e^{-2\pi i x(a+bi)}dx \leqslant Ce^{A|a+bi|},$$

as the exponential converges to 0 as $b \to \infty$. Conversely, if f is the Fourier transform, let $f_\epsilon(x) = f(x)e^{-\epsilon|x|}$. If one can show that $\check{f}_\epsilon$ is supported on $[-A, A]$ and $f_\epsilon \to f$ in L^2, we are done by Plancherel. One may define a family of Fourier transforms Φ_α on rotated half-planes by angle $2\pi\alpha$ through the origin, which one can also show are all analytic continuations of the Fourier transform in the upper half-plane. Then, formally $\check{f}(x) = \Phi_0(ix) - \Phi_\pi(ix)$, which can be made precise by the ϵ-scaling argument and concludes from $\Phi_0(\epsilon + 2\pi i t) - \Phi_\pi(-\epsilon + 2\pi i t) \to 0$ by uniqueness of analytic continuation. ∎

Corollary 2.2.2. *As a direct corollary of this, we can conclude that the Fourier transform of a compactly supported continuous function is an analytic function decaying at infinity, and so is not compactly supported by the maximum modulus principle.*

Here is an important generalization of these results.

Theorem 2.2.3 (Schwartz-Paley-Wiener)**.** *An entire function u is the Fourier transform of a compactly supported distribution v supported on $B(0, A)$ iff $u \ll |z|^N e^{A|z|}$. Moreover, $u \ll |z|^{-N} e^{A|z|}$ for all $N \geqslant 0$ iff $v \in C_c^\infty$.*

We know that the Fourier transform of a Gaussian decays like another Gaussian. The question is, can we do better? Turns out, we cannot. This is a reflection of the so-called "uncertainty principle" of Fourier transforms.

Proposition 2.2.1 (Uncertainty principle)**.** *For $f \in L^2(\mathbb{R})$ differentiable, $\|f\|_2 = 1$,*

$$\|xf\|_2 \|\xi f\|_2 \geqslant \frac{1}{16\pi^2},$$

with equality obtained only if f and $\widehat{f}$ are Gaussians.

Proof. Integrating by parts,

$$1 = \int |f|^2 dx = -\int 2x\text{Re } (f)\overline{f'},$$

so

$$1 \leqslant 2\|xf\|_2^2 \|f'\|_2^2 = 4\pi \|xf\|_2^2 \|\xi \widehat{f}\|_2^2,$$

and equality holds whenever $xf = f'$, which defines a Gaussian. □

Finally, it is worth mentioning the notion of **Fourier multipliers/symbols.** Define the Fourier symbol S_T of an operator T to be

$$S_T\{f\} = (\mathcal{F}^{-1} T \mathcal{F})\{f\}$$

whenever this is well-defined.

2.3 Fourier Series

By Hölder, it suffices for $f \in L^1(\mathbb{T}^k)$ to have a well-defined Fourier series $\widehat{f}$. Moreover, by Hilbert space theory, one deduces that $\mathcal{F} : L^2(\mathbb{T}^n) \to l^2(\mathbb{T}^n)$ is a unitary isometric isomorphism, so one has convergence of the Fourier series in L^2. By Riesz-Thorin, we have $\mathcal{F} : L^p \to l^{p'}$ for $1 \leqslant p \leqslant 2$.

Definition 2.3.1. The partial sums of the Fourier series of f on $\mathbb{T}$ are given by $S_n f = D_n * f$, where

$$D_n(x) := \sum_{k=-n}^{n} e^{2\pi i k x} = \frac{\sin((n + \frac{1}{2})x}{\sin x}$$

is the one-dimensional **Dirichlet kernel.** On $\mathbb{T}^n$, the Dirichlet kernel is $D_n = \prod_{k=1}^{N} D_N(x_i)$.

The Dirichlet kernel is unbounded in L^1 and so is not particularly nice to deal with, so we introduce a smoothed version.

Definition 2.3.2. The **Fejer** kernel is defined as $K_N := \frac{1}{N} \sum_{n=1}^{n} D_n = \frac{(1-\cos nx)}{n(1-\cos x)}$.

Theorem 2.3.1. *The Fejer kernel is an approximation to the identity.*

Remark 2.3.1. The Fejer and Dirichlet kernels both converge as distributions to the tempered distribution known as the **Dirac comb** $\Phi = \sum_{n\in\mathbb{Z}} \delta(x-n)$. However, since the Fejer kernel is an approximation to the identity, we have $K_n f \to f$ for all $f \in L^p(\mathbb{T}^n)$.

Lemma 2.3.1. *$S_n f \to f$ in L^p iff $\sup_n \|S_n\|_p < \infty$, where $S_n : L^p \to L^p$.*

Proof. One direction is immediate from Banach-Steinhaus, and the other follows from

$$\|S_n f - f\|_p = \|S_n(f - K_n f) + K_n f - f\|_p \leqslant \sup_n(\|S_n\|_p + 1)\epsilon.$$

□

Proposition 2.3.1. *$S_n : L^p \to L^p$ are uniformly bounded in L^p iff $1 < p < \infty$.*

Corollary 2.3.1. *Since S_n is not uniformly bounded in L^1 or L^∞, we get that $S_n f \to f$ iff $1 < p < \infty$.*

Remark 2.3.2. By Baire Category, one sees that the set of functions that converge at a particular point is meager in L^1.

Note that one has the following bounds on the decay of certain Fourier coefficients.

(a) If f is absolutely continuous, $\widehat{f}(n) \ll \frac{1}{n}$.

(b) If f is a function with bounded variation K, $|\widehat{f}(n)| \leqslant \frac{K}{2\pi|n|}$.

(c) If $f \in C^{0,\alpha}$ with $|f|_{C^{0,\alpha}} = K$, $|\widehat{f}(n)| \leqslant \frac{K}{|n|^\alpha}$.

We now provide a list of results regarding different types of convergence of Fourier series. Note that the proofs of these results are quite technical and are therefore omitted.

Theorem 2.3.2. *(a) If $1 < p < \infty, f \in L^p$, then $S_n f \to f$ in L^p.*

(b) If f is of bounded variation, then $S_n f \to f$ pointwise, and if f is continuous, $S_n f \to f$ uniformly.

(c) If f is α-Hölder continuous for $\alpha < 1$, then $S_n f \to f$ uniformly.

(d) $K_n f \to f$ a.e., uniformly if f is continuous, and in L^p if $f \in L^p$.

*(e) **Carleson's Theorem:** For $p > 1$, if $f \in L^p$, $S_n f \to f$ a.e.*

Remark 2.3.3. (d) follows from the fact that the Fejer kernel is an approximation to the identity, and (a) follows by Riesz-Thorin.

One may also ask about singly/doubly periodic holomorphic functions.

Proposition 2.3.2. *An entire 1-periodic function has an absolutely convergent Fourier series expansion*

$$f(z) = \sum_{n=-\infty}^{\infty} a_n e^{2\pi i n z}$$

iff $\limsup_n |a_n|^{\frac{1}{n}} = 0$. Moreover, every bounded entire 1-periodic function on the upper half-plane has a Fourier series expansion with only positive terms iff $\limsup_n |a_n|^{\frac{1}{n}} \leqslant 1$.

Proof. Note that $f(z) = F(e^{2\pi i z})$ for some holomorphic function $F : \mathbb{C}\backslash\{0\} \to \mathbb{C}$. Thus, F has a Laurent series expansion, which gives the Fourier series for f. The converse follows by completing the argument in the opposite direction. For the half-plane, we note that $F : \mathbb{D}\backslash\{0\} \to \mathbb{C}$, so using Riemann's removable singularity theorem yields a bounded holomorphic function with a power series expansion, which gives the Fourier series for f. □

2.3.1 Exercises

Problem 2.3.1 (Fall 2020 Problem 6). Show that for all odd $f \in C^1[-1,1]$,

$$\|f\|_{L^2} \leqslant \|f'\|_{L^2}.$$

Proof. Consider f as a periodic function and consider its Fourier series. By Plancherel,

$$\|f\|_{L^2} = \|\widehat{f}\|_{l^2} \leqslant \|in\widehat{f}\|_{l^2} = \|f'\|_{L^2},$$

where the inequality holds since $|n\widehat{f}(n)| \geqslant |\widehat{f}(n)|$ for $n > 0$ and $\widehat{f}(0) = \int_{-1}^{1} f = 0$. □

Problem 2.3.2 (Spring 2015 Problem 4, Wiener's Tauberian Theorem). Let $f \in L^1(\mathbb{R})$. Show that the translates of f, $f(x-a)$, are dense in $L^1(\mathbb{R})$ iff $\widehat{f}(\xi) \neq 0$. Similarly, show that for $f \in L^2$, the translates are dense iff $\widehat{f}$ is nonzero a.e.

Proof. This is a well-known result known as Wiener's Tauberian Theorem. We first show it in L^2. By the properties of the Fourier transform, $\widehat{f(x-a)} = e^{2\pi i\xi a}\widehat{f}(\xi)$. Then, suppose that $g \in L^2$ is orthogonal to all translates of f, i.e.

$$\int f(x-a)\overline{g(x)} = 0.$$

By Parseval's Theorem, this equals

$$\int \widehat{f}(x-a)\overline{\widehat{g}(x)} = \int e^{2\pi i\xi a}\widehat{f}(\xi)\widehat{g}(\xi) = 0$$

for all $a \in \mathbb{R}$. In particular, this implies that

$$\mathcal{F}^{-1}[\widehat{f}\widehat{g}](a) = 0.$$

for all a, and is thus equal everywhere. Since the inverse Fourier transform is injective. This implies that $\widehat{f}\widehat{g} = 0$, and since $\widehat{f}$ is nonzero a.e., $\widehat{g} = 0$, i.e. $g = 0$ a.e. Conversely, suppose that $\widehat{f}$ vanishes on a positive finite measure set X. Since the Fourier transform is an isometry on L^2, note that the translates of f are dense in L^2 iff $e^{2\pi i\xi a}\widehat{f}(\xi)$ is dense in L^2. However, $\chi_X \in L^2$ is orthogonal to all functions of the form $e^{2\pi i\xi a}$, which is a contradiction.

We now prove the more difficult version of this theorem. Suppose $\widehat{f}(\xi_0) = 0$ for some $\xi_0 \in \mathbb{R}$. Then, $\widehat{f}_a(\xi_0) = 0$ for all $a \in \mathbb{R}$, so the Fourier transform at ξ_0 vanishes for all functions in the span. However, the Fourier transform of a gaussian is everywhere nonzero, which is a contradiction. The other direction is *complicated* and over 100 pages in length. □

Problem 2.3.3 (Fall 2014 Problem 4). Define X as the set of $f \in L^2([0,\pi])$ that admit a representation of the form

$$f(x) = \sum_{n=0}^{\infty} c_n \cos(nx), \quad \|\langle n\rangle c_n\|_2 < \infty.$$

Show that if $f, g \in X$, then $fg \in X$.

Proof. By Fourier series, note that X is the set of $f \in L^2$ such that $\widehat{f}(n) = \widehat{f}(-n)$. and $\|\langle n\rangle \widehat{f}\|_2 < \infty$. First, if $f, g \in X$, then

$$\widehat{fg}(n) = \widehat{f} * \widehat{g}(n) = \sum_{k=-\infty}^{\infty} \widehat{f}(n-k)\overline{\widehat{g}(k)} = \sum_{k=-\infty}^{\infty} \widehat{f}(k-n)\overline{\widehat{g}(-k)} = \sum_{k=-\infty}^{\infty} \widehat{f}(-k-n)\overline{\widehat{g}(k)} = \widehat{fg}(-n).$$

Moreover, $\|\langle n\rangle \widehat{fg}\|_2 \leqslant \|\langle n\rangle \widehat{f} * \widehat{g}\|_2 \leqslant \|\langle n\rangle \widehat{f}\|_2 \|\widehat{g}\|_2 < \infty$, since $\widehat{g} \in l^2$ for $g \in L^2$. □

2.4 Convolutions

Recall the definition of a convolution:

Definition 2.4.1. The **convolution** of f and g is defined as

$$f * g(x) = \int_{\mathbb{R}} f(x-y)g(y)dy.$$

Here are some important properties of convolutions:

(a) $f * g = g * f$.

(b) $(f') * g = f * (g') = (f * g)'$.

(c) if f is C^k, $f * g$ is C^k.

(d) $\|f * g\|_1 = \|f\|_1 \|g\|_1$.

These properties provide for the following nice applications:

Theorem 2.4.1 (Approximation to the Identity). *Let $f \in L^p$. Then, if $\phi \in C_c^\infty(\mathbb{R}^n)$, $\|\phi\|_1 = 1$, is such that $\phi_\epsilon := \epsilon^{-n}\phi(\frac{x}{\epsilon}) \to \delta$ as $\epsilon \to 0$ (in the sense of distributions), then $f * \phi_\epsilon$ is smooth,* $\lim_{\epsilon\to 0} f * \phi_\epsilon = f$ *a.e., normally if f is continuous, and in L^p_{loc} if $f \in L^p_{loc}$.*

Proof. A.e. convergence follows from Radon-Nikodym and approximating by simple functions. The other types of convergence follow from the continuity of translation operators on L^p. □

The convolution of two functions measures their magnitude of intersection and has the following nice properties:

Lemma 2.4.1 (Steinhaus Theorem). *if $\mu(A) > 0$, $A - A$ contains an open neighborhood of 0.*

Proof. We in fact prove a stronger claim: if A, B are distinct sets of positive measure, there exists an x such that $(x - A) - B$ contains an open neighborhood of 0.

Consider the convolution of functions $\chi_A * \chi_B$. Note that

$$\|\chi_A * \chi_B\|_1 = \int\int |\chi_A(x-y)\chi_B(y)| dxdy = \|\chi_A\|_1 \|\chi_B\|_1 = \mu(A)\mu(B) > 0,$$

and since $\chi_A * \chi_B$ is continuous, there is an x such that on an open neighbor hood of x, $\chi_A * \chi_B(x) > 0$. But this precisely implies that $B \cap (y - A) \neq \varnothing$ for $y \in (x - \delta, x + \delta)$, i.e. $\delta' \in (x - A) - B$ for $|\delta'| < \delta$. □

Corollary 2.4.1 (Young's Convolution Inequality)**.** *Note that for an arbitrary $g \in L^p$, convolution with g defines a bounded operator $T : L^1 \to L^p$ and $L^q \to L^\infty$, since by Minkowski,*

$$\|Tf\|_p = \|f * g\|_p = \left\| \int f(y)g(x-y)dy \right\|_{L^p(dx)} \leqslant \int f(y)\|g\|_p dy \leqslant \|f\|_1 \|g\|_p$$

and

$$\|Tf\|_\infty = \sup \int f(y)g(x-y)dy \leqslant \|f\|_q \|g\|_p.$$

Thus, Riesz-Thorin guarantees that T is bounded as an operator from $L^r \to L^s$, i.e.

$$\|f * g\|_s \leqslant \|f\|_r \|g\|_p$$

for

$$\frac{1}{s} = 1 - t + \frac{t}{q}, \frac{1}{r} = \frac{1-t}{p} \implies \frac{1}{r} + \frac{1}{p} = 1 + \frac{1}{s}.$$

2.5 Layer Cake and Fubini

Often times, one wants to consider a different integration variable.

Lemma 2.5.1 (Chebyshev's Inequality)**.** *For $f \in L^p$,*

$$\mu\{x : |f(x)| > \lambda\} \leqslant \frac{\|f\|_p^p}{\lambda^p}$$

Proof.

$$\|f\|_p^p \geqslant \int_{|f|>\lambda} |f|^p \geqslant \mu(\{x : |f(x)| > \lambda\})\lambda^p.$$

□

Lemma 2.5.2 (Layer Cake Decomposition)**.** *For $f \in L^p$,*

$$\|f\|_p^p = \int_X |f|^p dx = \int_0^\infty p\lambda^{p-1}\mu(\{x : |f(x)| > \lambda\})d\lambda.$$

Proof. By Fubini,

$$\begin{aligned} \int_X |f|^p dx &= \int_X \int_0^{|f|} p\lambda^{p-1} d\lambda = \int_X \int_0^\infty p\lambda^{p-1} \chi_{\lambda<|f(x)|}(\lambda) d\lambda dx \\ &= \int_0^\infty p\lambda^{p-1} \int_X \chi_{|f(x)|>\lambda}(x) dx d\lambda = \int_0^\infty p\lambda^{p-1}\mu(\{x : |f(x)| > \lambda\})d\lambda. \end{aligned} \tag{1}$$

Note that if $f \in L^{p,w}$, then one ends up integrating $\frac{1}{x}$, which is almost in L^1. □

Remark 2.5.1. Intuitively, this states that the integral of a function can be approximated by horizontal rectangles lying below the graph of the function of width $\Delta\lambda$ and height $\mu\{x : |f| > \lambda\}$ at λ.

Remark 2.5.2. For $f \in L^p$, the function $\lambda \to \mu\{x : |f| > \lambda\}$ is called a **distribution function.** Namely, if μ is a probability measure and f is a random variable, then this precisely corresponds to the definition of a cumulative distribution function (cdf) in probability theory, equivalently defining the pushforward measure ν on $\mathbb{R}$ by $E \to \mu(f^{-1}(E))$. Then, the Radon-Nikodym derivative $\frac{d\nu}{d\mu}$ is the probability density function (pdf) of f.

Definition 2.5.1. Define the **symmetric decreasing rearrangement** A^* of a finite measure $A \subset \mathbb{R}^n$ to be the ball in $\mathbb{R}^n$ with the same measure as A. Given $f \geqslant 0 \in L^p$, the **symmetric decreasing rearrangement of** f is the unique positive radial function $f^*(r)$ such that

$$f^*(r) = \int_0^\infty \chi_{\{x:|f|>\lambda\}^*}(r)d\lambda.$$

Note that $\chi_{\{x:|f|>\lambda\}^*}(r) = 1$ iff $r \in \{x : |f| > \lambda\}^*$, i.e. $f^*(r)$ is the largest height λ of f for which the radius of $\{x : |f| > \lambda\}^*$ is greater than or equal to r.

Remark 2.5.3. Intuitively, one can think of slicing the peaks of the function f and putting them into the center, so that the value of $f^*(r)$ is the value of λ at which the volume of the peaks above λ of f exceeds the volume of the ball of radius r.

Remark 2.5.4. The defining quality of the symmetric decreasing rearrangement is that $\mu\{x : f > \lambda\} = \mu\{x : f^* > \lambda\}$.

Lemma 2.5.3. $\|f^*\|_p = \|f\|_p$.

Proof.

$$\|f\|_p^p = \int_0^\infty p\lambda^{p-1}\mu\{x : f > \lambda\}d\lambda = \int_0^\infty p\lambda^{p-1}\mu\{x : f^* > \lambda\} = \|f^*\|_p^p.$$

□

2.5.1 Exercises

Problem 2.5.1 (Fall 2010 Problem 4). Let $T : C_c(\mathbb{R}) \to C_c(\mathbb{R})$ be a linear transformation such that

$$\|Tf\|_\infty \leqslant \|f\|_\infty, \quad \mu\{x : |Tf(x)| > \lambda\} \leqslant \frac{\|f\|_1}{\lambda}.$$

Show that $\|Tf\|_2 \lesssim \|f\|_2$.

Proof. This is a consequence of the Marcinkiewicz interpolation theorem, and we reproduce a sample proof below. For $f \in C_c(\mathbb{R})$, write $f = g + h$, where $g = f\chi_{|f|<\frac{\lambda}{2}} + \frac{\lambda}{2}\chi_{|f|\geqslant\frac{\lambda}{2}}$. Then,

$$\{x : |f| > \lambda\} \subset \{x : |h| > \frac{\lambda}{2}\} \cup \{x : |g| > \frac{\lambda}{2}\},$$

where the latter set is empty. Then, by the first bound, $\mu\{x : |Tf| > \lambda\} \leqslant \mu\{x : |Th| > \frac{\lambda}{2}\}$. Writing the layer-cake decomposition, we now have

$$\begin{aligned}
\|Tf\|_p^p &= \int_0^\infty p\lambda^{p-1}\mu\{x : |Tf| > \lambda\}d\lambda \\
&\leqslant \int_0^\infty p\lambda^{p-1}\mu\{x : |Th| > \frac{\lambda}{2}\}d\lambda \\
&\leqslant \int_0^\infty p\lambda^{p-2}\int |h|dyd\lambda \\
&\leqslant \int_0^\infty \int p\lambda^{p-2}|f - f\chi_{|f|<\frac{\lambda}{2}} - \frac{\lambda}{2}\chi_{|f|>\frac{\lambda}{2}}|dxd\lambda \\
&\leqslant \int\int_0^{2|f|} p\lambda^{p-2}|f - \frac{\lambda}{2}|d\lambda dx \\
&\int |f|^{p-1}|f| + |f|^p dx \lesssim \|f\|_p^p.
\end{aligned}$$

□

Problem 2.5.2 (Fall 2020 Problem 5). Suppose $f \in L^1$ is such that $\int_E |f| \leqslant \sqrt{|E|}$ for all Borel $E \subset [0, 1]$. Show that $f \in L^p$ for $1 < p < 2$, but not necessarily in L^2.

Proof. Note that $\|f\|_1 \leqslant 1$, and moreover,

$$|\{x : |f| > \lambda\}|\lambda \leqslant \int_{|f|>\lambda} |f| \leqslant \sqrt{|\{x : |f| > \lambda\}|},$$

i.e.

$$|\{x : |f| > \lambda\}| \leqslant \frac{1}{\lambda^2}.$$

Then, by the layer-cake decomposition, for $p > 2$,

$$\|f\|_p^p \leqslant \int_{|f|\leqslant 1} |f|^p dx + \int_1^\infty p\lambda^{p-1}|x : |f| > \lambda|d\lambda \leqslant 1 + \int_1^\infty p\lambda^{p-3} < \infty$$

whenever $p < 2$. Moreover, $\frac{1}{\sqrt{x}}$ is in $L^1([0, 1])$ but not $L^2([0, 1])$, and since $\sqrt{x} - \sqrt{y} \leqslant \sqrt{x - y}$ for all $x \geqslant y$, the inequality holds on open intervals, and therefore on all open sets, so by regularity of the Lebesgue measure on $[0, 1]$, it holds for all Borel sets $E \subset [0, 1]$. □

2.6 Density Arguments

Density arguments typically rely on one of the following theorems or statements:

(a) Stone-Weierstrass: A *-subalgebra of $C(X)$ for compact Hausdorff X that separates points and does not vanish at any point is dense in $C(X)$.

(b) If μ is Borel, characteristic functions of open intervals are dense in characteristic functions of measurable sets, and the **span** of characteristic functions (of measurable sets) are dense in $L^p, 1 \leqslant p \leqslant \infty$.

(c) If X is an LCH space and μ is a Radon measure, $C_c(X)$ is dense in $L^p(X, \mu)$ for $1 \leqslant p < \infty$. In particular, this holds for every locally finite measure on $\mathbb{R}^n$.

2.6.1 Exercises

Problem 2.6.1 (Spring 2020 Problem 4). Show $\sin(x^n) \overset{*}{\rightharpoonup} 0$ in $L^\infty([0,2])$.

Proof. Note that $\sin(x^n) \to 0$ on $[0,1)$, so by dominated convergence, $\int_0^1 f \sin(x^n) \to 0$. For the interval $[1,2]$, we appeal to a density argument, showing that the statement is true whenever f is the characterstic function of a closed interval. Indeed, for $1 \leqslant a < b \leqslant 2$,

$$\left|\int_a^b \sin(x^n) dx\right| = \left|\int_{a^n}^{b^n} \frac{y^{\frac{1}{n}}}{n} \sin(y) dy\right| \leqslant \frac{1}{(1-n)b^{n-1}} - \frac{1}{(1-n)a^{n-1}} \to 0$$

as $n \to \infty$, so by density, the argument is complete. ☐

Problem 2.6.2 (Spring 2020 Problem 1). Suppose $f \in C_c^\infty$ satisfies $\int_{\mathbb{R}} f(x) e^{-tx^2} dx = 0$, for any $t \geqslant 0$. Show that f is odd.

Proof. We first reduce the problem by defining the even function $g(x) = f(x) + f(-x)$ and showing that g is identically zero, given that $\int_{\mathbb{R}} g(x) e^{-tx^2} = 0$ for all $t \geqslant 0$. By symmetry, this implies that $\int_0^\infty g(x) e^{-tx^2} = 0$ for all $t \geqslant 0$. Suppose f is supported on $[-R, R]$. Note that the algebra generated by the functions $\{e^{-tx^2} : t \geqslant 0\}$ on $[0, R]$ is a unital algebra that separates points, so by Stone-Weierstrass, it is dense in $C([0, R])$ in the uniform norm. In particular, one may take an element a in the algebra such that $\|a - g\|_\infty < \epsilon$. Note that the assumptions of the problem imply that $\int ga = 0$, so

$$\int_0^\infty g^2 dx \leqslant R\epsilon \|g\|_1 \to 0$$

as $\epsilon \to 0$. Thus, g is identically zero, i.e. f is odd. ☐

2.7 Oscillatory Integrals

Many times in harmonic analysis, one aims to asymptotically estimate the magnitude of an integral of the form

$$\int a(x) e^{i\lambda\phi(x)} dx.$$

The theory of oscillatory integrals and the method of stationary phase are powerful tools for estimating such integrals. First, one has the trivial bound

$$I(\lambda) := \left|\int_J e^{i\lambda\phi(x)} dx\right| \leqslant \mu(J).$$

This bound is achieved iff ϕ is constant, so the decay of this integral is linked to the noncostancy of ϕ. One way to achieve this is to require $|\phi'| \geqslant c > 0$. However, that turns out to be not enough. One additional assumption, for instance, is monotonicity.

Lemma 2.7.1 (Van der Corput). *If $\phi : \mathbb{R} \to \mathbb{R}$ is smooth, $|\phi'| \geqslant c > 0$, and ϕ' is monotonic, then* $|I(\lambda)| \ll \frac{1}{c\lambda}$.

Proof. Integrating by parts and using fundamental theorem of calculus on the second integral,

$$I(\lambda) = \int_a^b \frac{1}{i\lambda\phi'(x)} \frac{d}{dx} e^{i\lambda\phi(x)} dx = \left[\frac{1}{i\lambda\phi'(x)} e^{i\lambda\phi(x)}\right]_a^b - \frac{1}{i\lambda} \int_a^b \frac{d}{dx}\left[\frac{1}{\phi'(x)}\right] e^{i\lambda\phi(x)} dx = O(\frac{2}{c\lambda}) + O(\frac{2}{\lambda c}) = O(\frac{1}{\lambda c}).$$

□

3 Functional Analysis

3.1 Hilbert Space Theory

The following are main theorems and lemmas to be used from the theory of Hilbert spaces:

(a) Every Hilbert space admits an orthonormal basis e_n.

(b) **Parseval's Identity:** The orthonormal basis satisfies

$$\left\| \sum_{n=1}^{\infty} a_n e_n \right\| = \|(a_n)\|_{l^2}.$$

(c) For every closed convex subset W and any vector $v \notin W$, there exists a unique $w \in W$ such that $\|v - w\| = \inf_{w' \in W} \|v - w'\|$.

(d) For every closed subspace W of V, there exists an orthogonal decomposition of V as $V = W \oplus W^\perp$.

(e) **Riesz Representation Theorem:** For any $\phi \in V^*$, there exists a unique $v \in V$ such that $\phi(w) = \langle w, v \rangle$.

Proof. We prove the Riesz representation theorem. Let $\phi \in V^*$ and consider the decomposition $V = \ker\phi \oplus \ker\phi^\perp$. Pick $x_0 \in \ker\phi^\perp$ and notice that

$$\phi\left(x - \frac{\phi(x)}{\phi(x_0)} x_0\right) = 0.$$

This implies that

$$\left\langle x - \frac{\phi(x)}{\phi(x_0)} x_0, x_0 \right\rangle = 0 \implies \phi(x) = \langle x, \frac{\phi(x_0)}{\|x_0\|^2} x_0 \rangle.$$

Then, uniqueness is easily checked. □

3.1.1 Exercises

Problem 3.1.1 (Fall 2009 Problem 1). Find a closed subset in $L^2([0,1])$ with no element of smallest norm.

Proof. Let $X = \{f_n\}$, where $f_n = \sqrt{n+1}\chi_{[0,\frac{1}{n}]}$. Then, $\|f_n\|_2^2 = \frac{n+1}{n}$, and any subset of f_n converges to 0 a.e. Then, if some subsequence satisfied $f_n \to f$ in L^2 with $\|f\|_2 \neq 0$, it would have a subsequence that converges a.e., so the subsequence would have to converge to 0. But that is a contradiction, since $\|0\|_2 \neq 1$. Thus, X is a closed nonempty subset of L^2 with no element of smallest norm. $\square$

Problem 3.1.2 (Fall 2009 Problem 7). Define a unitary operator on a complex Hilbert space, and show that if S is unitary, then $S - \lambda I$ is invertible for $|\lambda| < 1$. Finally, show that if one defines

$$h(\lambda) = \langle (S + \lambda I)(S - \lambda I)^{-1} v, v \rangle,$$

then Re h is a positive harmonic function.

Proof. A unitary operator S is one that satisfies $\langle Sv, Sw \rangle = \langle v, w \rangle$ for all $v, w \in V$, or equivalently, one such that $SS^* = S^*S = I$. In particular, $\|SS^*\| \leqslant \|S\|^2 = 1$, so $\|S\| \leqslant 1$. Clearly, if $\lambda = 0$, $S - \lambda I$ is invertible with inverse S^*. If $\lambda \neq 0$, S is unitary and $\frac{1}{|\lambda|} > \|S\| = \|S^*\|$, so I claim that

$$(S - \lambda I)^{-1} = \frac{1}{S - \lambda I} = S^{-1} \frac{1}{I - \lambda S^{-1}} = S^* \sum_{n=0}^{\infty} (\lambda S^*)^n$$

is the inverse of $S - \lambda I$. Indeed it is a well-defined operator, as the series converges absolutely, since $\|(\lambda S^*)^n\| \leqslant |\lambda|^n \|S^*\|^n \leqslant (|\lambda| \|S\|)^n$, which is a geometric series that converges, and one can formally multiply the series with $S - \lambda I$ to check that it yields the identity.
Finally, define h as above. $\square$

Problem 3.1.3 (Fall 2010 Problem 6). Let V be the Hilbert space of holomorphic functions $f : \overline{\mathbb{D}} \to \mathbb{C}$ such that $f(z) = \sum_n a_n z^n$ and $\|f\| = \|\langle n \rangle \widehat{f}\|_2 < \infty$, where $\langle n \rangle = \sqrt{1 + n^2}$. Show that $L : f \to f(1)$ is a bounded linear functional on V, find the element g that represents L, and $f \to \text{Re } L(f)$ achieves a unique maximum on the set $X = \{f : \|f\| \leqslant 1, f(0) = 0\}$ and find this maximum.

Proof. a) Clearly, L is a linear functional. Then, by Cauchy-Schwarz,

$$|f(1)| = \sum_n \langle n \rangle^{-1} (\langle n \rangle^2 \widehat{f}(n)) \leqslant \|\langle n \rangle^{-1}\|_2 \|f\| \lesssim \|f\|.$$

b) We want to find $g \in V$ such that

$$f(1) = \langle f, g \rangle = \sum_n \langle n \rangle^2 \widehat{f} \overline{\widehat{g}}.$$

In particular, note that

$$\sum_n a_n = f(1) = \sum_n a_n \langle n \rangle^2 \overline{\widehat{g}},$$

so setting $\widehat{g}(n) = \langle n \rangle^{-2}$, we get that $g = \sum_n \widehat{g}(n) z^n \in V$ represents L. Note that $g \in V$ by direct computation.
c) To show that Re$(f(1))$ achieves its maximal value on X, we use the representation of the linear functional. If $\|f\| \leqslant 1$ and $f(0) = \widehat{f}(0) = 0$, then Re$f(1) \leqslant |f(1)| \leqslant \|\langle n \rangle^{-1}_{n \geqslant 1}\|_2$. Let Let $f = g - 1$. Then, $|f(1)| = \|f\|$, $f(0) = 0$, $\|f\| = \sqrt{\|g\|^2 - 1} \leqslant 1$. One may use Cauchy-Schwarz to show uniqueness, thus completing the proof. $\square$

Problem 3.1.4. Let $E \subset L^2([0,1])$ be a closed subset such that $E \subset C([0,1])$. Show that E is finite dimensional.

Proof. The proof roughly follows in 4 steps. First, notice that by Hölder, $\|f\|_2 \leqslant \|f\|_\infty$ for all $f \in E$. Next, consider the inclusion $(E, \|\cdot\|_2) \subset C([0,1])$. If $f_n \to f$ in L^2 and $f_n \to g$ in L^∞, then $f = g$ a.e., so by the closed graph theorem, the inclusion is continuous, i.e. $\|f\|_\infty \leqslant C\|f\|_2$. Now, for any $f \in E$, evaluation at $x \in [0,1]$ is a continuous linear functional, so by Riesz representation on the Hilbert space E, for some $g_x \in E$, $f(x) = \langle f, g_x \rangle$ and so $g_x(x) = \|g_x\|_2^2 \leqslant C\|g_x\|_2$, i.e. $\|g_x\|_2 \leqslant C$. Then, for any orthonormal basis f_i of E, by Bessel's,

$$\sum_i |f_i(x)|^2 = \|g_x\|_2^2 \leqslant C^2,$$

so integrating on both sides yields $|I| \leqslant C^2$, i.e. E is finite-dimensional. □

Here are some important results and problems from functional analysis.

Theorem 3.1.1 (Hahn-Banach). *Let V be a normed vector space and $W \subset V$ be a subspace. If $\phi : W \to \mathbb{C}$ is a linear map (not necessarily bounded) that is bounded by a seminorm $p : V \to \mathbb{R}$ on W, then ϕ extends to a map $\Phi : V \to \mathbb{C}$ bounded by p on V.*

Corollary 3.1.1. *(a) If $W \subset V$ is a closed subspace and $x \in V \backslash W$, there exists $\phi : V \to \mathbb{C}$ that vanishes on W, $\|\phi\| = 1$, and $\phi(x) = 1$.*

(b) Every continuous functional $W \subset V$ extends to a continuous functional of the same norm on V.

*(c) **Geometric Hahn-Banach:** If A, B are two closed convex disjoint nonempty subsets of V, then there exists a linear functional $\phi : V \to \mathbb{R}$ and some $c \in \mathbb{R}$ such that $\sup_A \phi(x) < c < \inf_B \phi(x)$, i.e. A, B are separated by the hyperplane $\phi^{-1}(c)$.*

(d) The map $i : X \to X^{**}$ is injective and isometric.

Theorem 3.1.2 (Open Mapping Theorem). *Let $T : X \to Y$ be a continuous linear map between Banach spaces. Then, either T is surjective and open, or the image of T is a set of the first category in V.*

Theorem 3.1.3 (Closed Graph Theorem). *If $T : X \to Y$ is a map between Banach spaces then if $\{(x, Tx)\} \subset X \times Y$ is closed, then T is continuous.*

Theorem 3.1.4 (Uniform Boundedness Principle). *Consider a family of bounded operators $T_\alpha : X \to Y$ between Banach spaces such that for each x, $\|T_\alpha x\| \leqslant C_x$ for all α for some constant depending on x. Then, $\|T_\alpha\| \leqslant C$ for some C for all α.*

Proof. Consider the sets $X_n = \{x : \sup_\alpha \|T_\alpha x\| \leqslant n\}$. By the Baire category theorem, one of these sets X_n contains an open ball $B(x_0, \epsilon)$. Then,

$$\sup_{\|u\| \leqslant 1} \|T_\alpha u\| = \sup_{\|u\| \leqslant 1} \left\| \frac{T_\alpha(x_0 + \epsilon u) - T_\alpha x_0}{\epsilon} \right\| \leqslant \frac{2n}{\epsilon}.$$

□

Theorem 3.1.5 (Banach-Alaoglu)**.** *If B is a Banach space, then the unit ball in B^* is weak-* compact.*

Proof. Define $B_x := \{z \in \mathbb{C} : |z| \leqslant \|x\|\}$, and consider $A := \prod_{x \in B(0,1)} B_x$, which is compact as a product of compact spaces by Tychonov's theorem. Then, if B^* is the unit ball in the weak$-^*$ topology, the map $\Phi : B^* \to A$ given by $\Phi_x(\phi) = \phi(x)$ is a homeomorphism onto a subset of A. By Hahn-Banach, the map is injective, and it it clearly continuous with respect to the weak$-^*$ topology. Finally, it easily checked that the image is closed in A and $f_n \to f$ weakly iff $\Phi(f_n) \to \Phi(f)$, so the map is a homeomorphism. Thus, B^* is homeomorphic to a compact set and is therefore compact. □

Often times, one wants to show certain types of compactness/weak compactness. Here we provide an overview of the conditions necessary to obtain such result, namely, considering the conditions of separability and reflexivity.

Definition 3.1.1. B is **reflexive** if the isometric embedding into the second dual $i : B \to B^{**}$ is a Banach space isomorphism, i.e. the weak and weak-* topologies on B^* coincide. B is **separable** if it has a countable dense subset.

Remark 3.1.1. The following is an extremely important remark: **since the weak/weak* topology is not necessarily metrizable, weak compactness and weak sequential compactness are NOT EQUIVALENT.**

Remark 3.1.2. Moreover, **it is extremely important to distinguish the metrizability of the entire space versus a compact set.** For instance, we will show that the weak/weak* topology is never metrizable on a space X, but the weak-* topology **on the unit ball** is metrizable if X is separable. In particular, this implies that in general, **if S is weakly sequentially closed, S is not necessarily weakly closed.** However, this is true, for example, on weakly bounded sets, since in that case the topology is metrizable, or for separable reflexive spaces.

Remark 3.1.3. We will use the letter X to denote a general vector space and B to denote a Banach space.

Lemma 3.1.1. *(a) Y is Banach iff $\mathcal{B}(X, Y)$ is Banach.*

(b) If X^* is separable, then X is separable.

(c) X is reflexive iff X^* is reflexive.

(d) X is reflexive and separable iff X^* is reflexive and separable.

(e) A Hilbert space is reflexive.

(f) X is separable iff the weak- topology on the unit ball of X^* is metrizable.*

(g) X^* is separable iff the weak topology on the unit ball of X is metrizable.

Proof. We prove (b) and (e). For (b), pick a dense subset ϕ_n of the unit sphere in B^*, and pick a sequence x_n on the unit sphere of X such that $\phi(x_n) \geqslant \frac{1}{2}$. Suppose that the $\mathbb{Q}$-span of x_n is not dense in X. Then, by Hahn-Banach, there exists a nonzero linear functional $\psi \in X^*$ with $\|\psi\| = 1$ that vanishes on the $\mathbb{Q}$-span of x_n. But then, for any ϕ_n,

$$|\psi(x_n) - \phi_n(x_n)| \geqslant \frac{1}{2},$$

contradicting the density of ϕ_n.
For (f), if x_n is a dense countable subset of the unit sphere, it suffices to define the metric

$$\rho(\phi, \psi) = \sum_{n=0}^{\infty} 2^{-n} \frac{(\phi - \psi)(x_n)}{1 + (\phi - \psi)(x_n)}.$$

It is then easy to see that $\phi_n \xrightarrow{*} \phi$ iff $\phi_n(x_m) \to \phi(x_m)$ for all m, i.e. this metric defines the weak-* topology. (g) follows similarly.

□

Lemma 3.1.2. *The weak/weak–* topology is never metrizable.*

Proof. Suppose d is a metric for the topology, consider the $U_n = \{x : d(x, 0) < \frac{1}{n}\}$, these are weakly open and therefore unbounded. But if $x_n \in U_n$, $\|x_n\| \geqslant n$, $x_n \rightharpoonup 0$, so x_n is bounded, a contradiction. □

Lemma 3.1.3. *In an infinite-dimensional normed vector space, the weak closure of the unit sphere is the unit ball.*

Proof. One inclusion is clear - since a convex set is closed if and only if it is weakly closed, B is weakly closed, and so $\overline{S^w} \subseteq B$. Conversely, recall that the weak topology is the coarsest topology on H, such that the linear functional evaluation maps $x \to \phi(x) = \langle x, \phi \rangle$ are continuous. Thus, the basic open neighborhoods in the weak topology of some $x \in H$ are the sets $U = \{y \in H : \langle y - x, \phi_i \rangle < \epsilon, i = 1, ..., n\}$. For $x \in B$, note that $y - x \in \bigcap_{i=1}^{n} \ker \phi_i$, and since kernels of linear functionals have finite codimension and X is infinite-dimensional, the intersection of the kernels is infinite-dimensional and therefore contains a line $L = \{tv : t \in \mathbb{R}, v \in X\}$ through the origin. In particular, since $\|x\| \leqslant 1$, if $y - x \in L$, so $y = x + L$ intersects S (since for $t = 0$ one has $\|y\| \leqslant 1$ and for t large $\|y\| \to \infty$.) Thus, any basic open neighborhood of $x \in B$ intersects S, i.e. S is weakly dense in B. Along with the other inclusion, it follows that $\overline{S^w} = B$. □

Lemma 3.1.4. *If X^* is separable, then there exists a sequence $x_n \in X, \|x_n\| = 1$, such that $x_n \rightharpoonup x$ for any $\|x\| \leqslant 1$.*

Proof. Since X^* is separable, the unit ball in X is weakly metrizable, so sequential weak closedness agrees with weak closedness. □

Remark 3.1.4. It follows that the unit sphere is never weakly closed. If X^* is separable, the above lemma shows that it is also not weakly sequentially closed. However, if $X = l^1, X^* = l^\infty$, which is not separable, and since l^1 has the Schur property, weak sequential convergence and norm convergence are equivalent, so the unit sphere is weakly sequentially closed but not weakly closed.

Lemma 3.1.5. *Given linearly independent functionals $\phi_i \in X^*, i = 1, ..., n$ of norm 1, and $|c_i| \leqslant 1, i = 1, ..., n$, there exists $x \in X$ with $\|x\| \leqslant 1$ s.t. $\phi_i(x) = c_i$. In finite dimensions this becomes a simple matrix problem.*

Proof. Note that $\phi_i(x) = c_i$ for $c_i \neq 0$ is equivalent to $(\phi_i - \frac{c_i}{c_j}\phi_j)(x) = 0, i \neq j$ assuming that $\phi_1(x) = c_1$. Clearly, such an x exists, as it is in the kernel of finitely many linear functionals in an infinite dimensional space and can be scaled appropriately to satisfy $\phi_1(x) = c_1$. □

Theorem 3.1.6 (Goldstine's Theorem)**.** *The image of the unit ball under the embedding $i : X \to X^{**}$ is weak-* dense in the unit ball of X^{**}.*

Proof. Let $y \in B_{X^{**}}$, so that $|y(\phi)| \leqslant \|\phi\|$. Notice that $\{x \in X^{**} : |(y-x)(\phi_i)| < \epsilon, i = 1, ..., n\}$ is a basic weak-* neighborhood of y in $B_{X^{**}}$. Without loss of generality, one may take $\{\phi_i\}$ to be linearly independent. But by the above lemma, one can find $x \in B_X$ such that $\phi_i(x) = y(\phi_i), i = 1, ..., n$ so $i(B_X)$ intersects every open weak-* neighborhood of y. ☐

Theorem 3.1.7. *The following are equivalent:*

(a) B is reflexive.

(b) ***Kakutani's Theorem:*** *The unit ball in B is weakly compact.*

(c) ***Eberlein-Smulian Theorem:*** *The unit ball in B is weakly sequentially compact.*

Proof. If B is reflexive, then by Banach-Alaouglu, the unit ball in B is weak-* compact and therefore weakly compact. Conversely, the image of the unit ball under the isometric embedding is weak-* closed and and dense in the unit ball of B^{**}, so it comprises the entire unit ball, i.e. $i : B \to B^{**}$ is bijective and thus B is reflexive. ☐

Corollary 3.1.2. *(a) If B is separable or reflexive, then the unit ball in B^* is weak—* sequentially compact.*

(b) If B is both reflexive and separable, all unit balls are compact/sequentially compact in all weak topologies.

Corollary 3.1.3. *A reflexive Banach space is* ***weakly sequentially complete****, i.e. every weak-* Cauchy sequence converges.*

3.2 Unbounded Operators and Adjoints

Definition 3.2.1. An **unbounded** operator $T : D(T) \subset X \to Y$ is a linear map defined on a subspace $D(T)$, called the **domain of** T. If $D(T)$ is dense in X, T is said to be **densely defined.** T is said to be **closed** if its graph $\{(x, Tx)\} \subset X \times Y$ is closed.

Proposition 3.2.1. *$T : D(T) \subset X \to Y$ is bounded on $D(T)$ iff $D(T)$ is closed and T is closed.*

Proof. A bounded operator is clearly closed and $D(T)$ is closed. The converse follows from the closed graph theorem. ☐

Corollary 3.2.1. *Thus, closed unbounded operators are never defined on X. One typically works with densely defined closed unbounded operators.*

We now establish a correspondence between graphs and unbounded operators.

Proposition 3.2.2. *There is a one-to-one correspondence*

$$\text{unbounded (closed) operators } T \text{ on } D(T) \longleftrightarrow \text{(closed) subspace } C \text{ of } X \times Y \text{ s.t. } ((0,y) \in C \implies y = 0), \pi(C) =$$

Proof. One direction is obvious. The other follows from defining $Tx = y$ for $(x, y) \in C$ and checking that this is indeed linear. ☐

Definition 3.2.2. A **closable operator** T is an operator such that the closure of its graph satisfies $(0, y) \in C \implies y = 0$. The **closure** $\overline{T}$ is the operator corresponding to the closure of the graph of T.

Remark 3.2.1. Note that the closure $\overline{T}$ of a closable operator is an extension of the corresponding graph. In general, there exists a (highly nonunique) extension of any operator. Additionally, the condition of being closed is weak in the following sense: if T is defined on a dense domain B and A is a dense subdomain, then the closure of $T|_A$ and $T|_B$ need not be equal (take, for instance, T to be identically 0 on A and nonzero somewhere on B). However, it turns out that this distinction disappears for self-adjoint operators.

Example 3.2.1. The derivative operator $\frac{d}{dx} : C^1([0,1]) \subset C([0,1]) \to C([0,1])$ is a closed, densely defined, unbounded operator. To see this, let $f_n \in C^1$ be s.t.

$$(f_n, f_n') \to (f, g) \in C([0,1]) \times C([0,1]).$$

Then, by a classic result on convergence of derivatives, it follows that $f \in C^1$ and $f' = g$. This implies that $C^1([0,1])$ is not closed in $C([0,1])$. However, if we replace the domain by $C^\infty([0,1])$, the operator is not closed, since there is no guarantee that $f \in C^\infty$.

Example 3.2.2. The unbounded densely defined operator $T : C([0,1]) \to L^2([0,1])$ given by $Tf = f(0)$ is not closable.

Definition 3.2.3. For a bounded operator $T : X \to Y$ between Banach spaces, define the **adjoint** $T^* : Y^* \to X^*$ by $T^*(\phi)(x) = \phi(Tx)$.

For an unbounded densely-defined operator $T : D(T) \subset X \to Y$ between Banach spaces, define

$$D(T^*) = \{y^* \in Y^* : \exists C \geqslant 0, |y^*(Tx)| \leqslant C\|x\|_X, x \in D(T)\}.$$

Then, one can uniquely define $T^* : D(T^*) \subset Y^* \to X^*$ by $T^*(\phi)(x) = \widehat{\phi}(Tx)$, which is the Hahn-Banach extension of $T^*(\phi)$ to all of X. T^* is called the **formal adjoint** of T.

Definition 3.2.4. An operator T is called **symmetric (or formally self-adjoint)** if T^* is an extension of T, and **self-adjoint** if $T^* = T$. If T is symmetric and its closure is self-adjoint, then T is called **essentially self-adjoint.**

Proposition 3.2.3. *(a) T^* is always closed.*

(b) T closable $\iff$ T^* densely-defined, in which case $\overline{T} = T^{**}$.

(c) (Hellinger-Toeplitz): A symmetric operator T with $D(T) = H$ is bounded.

Proof. If T^* is densely defined, one can easily check that T^{**} is the closure of T. Conversely, if T is closable. The other direction is slightly more complicated.

The first and last statement are a direct consequence of the closed graph theorem and the fact that if $(x_n, Tx_n) \to (x, y)$,

$$\langle z, Tx_n \rangle = \langle Tz, x_n \rangle \to \langle z, y \rangle = \langle Tz, x \rangle \implies \langle z, Tx - y \rangle = 0 \quad \forall z.$$

□

Corollary 3.2.2. *Since we can always pass from a closable operator to a closed operator, it follows that T is closed and densely-defined iff T^* is. Moreover, this implies that symmetric operators are closable and densely-defined.*

We have seen that closed extensions of operators need not be unique. What about symmetric extensions of symmetric operators?

Example 3.2.3. Take $T = -\partial_x^2$ on $L^2([a,b])$ with $D_T = \{f \in C^\infty : f^{(n)}(a) = f^{(n)}(b) = 0, n \geqslant 0\}$. It is easy to see that T is positive and symmetric. In particular, we can define the following extensions on larger domains:

$$T_{\alpha,\beta} = \{f \in C^\infty : f(a) = \alpha f(b), f'(a) = \beta f'(b)\}.$$

It turns out the extension is symmetric iff $\langle \alpha, \beta \rangle = 1$, and in fact, any two such extensions do not have a common extension. This example shows the importance of boundary conditions in these sorts of problems. However, there is a **unique self-adjoint** extension (which is given by the closure $\overline{T} = T^{**}$ of T), which occurs iff T is essentially self-adjoint.

Proposition 3.2.4. *If $T : X \to Y$ is a bounded linear operator between Hilbert spaces,*

$$\ker T^\perp = \overline{\operatorname{ran} T^*}, \operatorname{ran} T^\perp = \ker T^*.$$

Proof.

$$0 = \langle Tx, y \rangle = \langle x, T^* y \rangle.$$

□

Corollary 3.2.3. *T is injective iff T^* has dense image.*

Lemma 3.2.1. *A bounded linear map $T : X \to Y$ between Banach spaces is injective and has closed range iff T is bounded below.*

Proof. If T is bounded below, then $\|Tx\| \geqslant C\|x\|$, so $Tx = 0$ implies $x = 0$, i.e. T is injective. Similarly, if $Tx_n \to y$, then x_n is Cauchy, so $x_n \to x$, and $y = Tx$. Conversely, $T : X \to \operatorname{ran} T$ is an isomorphism, so by the open mapping theorem its inverse is bounded. □

Lemma 3.2.2. *If T is injective and has closed range, then T^* is surjective.*

Proof. $T : X \xrightarrow{\sim} \operatorname{ran} T \hookrightarrow Y$ is an isomorphism, which induces an isomorphism $T^* : Y^* \twoheadrightarrow (\operatorname{ran} T)^* \xrightarrow{\sim} X^*$. □

Theorem 3.2.1 (Closed Range Theorem)**.** *TFAE for a closed densely-defined operator $T : X \to Y$:*

(a) $\operatorname{ran} T$ *is closed.*

(b) $\operatorname{ran} T^*$ *is closed.*

(c) $\operatorname{ran} T = (\ker T^*)^\perp$.

(d) $\operatorname{ran} T^* = \ker T^\perp$.

Proof. We remark that

$$T : A \hookrightarrow B \to C \implies T^* : C^* \to B^* \twoheadrightarrow A^*,$$

and

$$T : B \twoheadrightarrow B/A \to C \implies T^* : C^* \to (B/A)^* \hookrightarrow B^*.$$

It suffices to prove that (a) implies (d). Since

$$T : X \twoheadrightarrow X/\ker T \xrightarrow{\sim} \operatorname{ran} T \hookrightarrow Y$$

is an isomorphism,

$$T^* : Y^* \twoheadrightarrow (\operatorname{ran} T)^* \xrightarrow{\sim} (X/\ker T)^* \cong \ker T^\perp \hookrightarrow X^*$$

is an isomorphism, so $\operatorname{ran} T^* = \ker T^\perp$. □

Remark 3.2.2. The rough conclusion of this section is that T surjective implies T^* injective, and T bounded below implies T^* surjective.

Definition 3.2.5. A subspace $W \subset V$ of a Banach space is said to be **complemented** if $V = W \oplus Z$ as Banach spaces, and the projections are continuous.

Remark 3.2.3. While it is true that $V = W \oplus Z$ as vector spaces (because of algebra), the additional requirement that the projections are continuous makes the statement deeper for Banach spaces. Note that in particular this implies that if $W \subset V$ is complemented, then $W \subset V$ is closed.

3.3 Spectral Theory

Definition 3.3.1. A Frechet-differentiable function $F : U \to Y$ between complex Banach spaces is said to be **holomorphic in** U.

Definition 3.3.2. For an unbounded linear operator $T : D(T) \subset X \to Y$, the **spectrum** $\sigma(T) \subset \mathbb{C}$ of T is the set of λ for which $(T - \lambda I)^{-1} : Y \to D(T)$ exists and is bounded. $\rho(T) = \mathbb{C}\backslash\sigma(T)$ is called the **resolvent set of** T, and $(T - \lambda I)^{-1}$ is known as the **resolvent operator of** T.

Remark 3.3.1. More generally, one may define the spectrum of an element a of a unital algebra A over a field $\mathbb{K}$ as the set of $\lambda \in \mathbb{K}$ s.t. $a - \lambda$ is not invertible.

Remark 3.3.2. By the closed graph theorem $\lambda \in \sigma(T)$ iff $T - \lambda I$ is not bijective.

Proposition 3.3.1. *(a) If T is not closed, $\sigma(T) = \mathbb{C}$. Otherwise, $\sigma(T) \subset \mathbb{C}$ is closed (possibly empty), and if T is bounded, $\sigma(T) \subset \overline{B(0, \|T\|)}$ is nonempty and compact.*

(b) $\lambda \to (T - \lambda I)^{-1}$ is holomorphic on $\rho(T)$.

(c) If T is invertible, $\lambda \in \sigma(T) \iff \lambda^{-1} \in \sigma(T^{-1})$.

Proof. Note that $(T - \lambda I)^{-1} = -\lambda^{-1}(I - \frac{T}{\lambda})^{-1}$, which has a geometric power series expansion and is therefore holomorphic whenever $\lambda > \|T\|$, so $\sigma(T)$ is bounded. Moreover, if $\lambda_0 \in \rho(T)$,

$$(T - \lambda)^{-1} = ((T - \lambda_0) - (\lambda - \lambda_0))^{-1} = (T - \lambda_0)^{-1}(I - (\lambda - \lambda_0)(T - \lambda_0)^{-1})^{-1},$$

which has a geometric power series expansion and is therefore holomorphic whenever $|\lambda - \lambda_0| < \frac{1}{\|(T-\lambda_0)^{-1}\|}$. Thus, $\sigma(T)$ is closed and bounded, i.e. compact. Finally, if $\sigma(T)$ is empty, then $\lambda \to (T - \lambda I)^{-1}$ defines a bounded entire function (since $\|(T - \lambda)^{-1}\| \leqslant \frac{1}{\|T-\lambda\|} \leqslant \frac{1}{|\|T\|-|\lambda||}$ for $|\lambda|$ large, which by Liouville's theorem implies that it must be constant, a contradiction. □

Remark 3.3.3. The version of Liouville's theorem used here is that a bounded entire function with values in a complex normed vector space is constant. This can be proven using the classical Liouville theorem and composing with bounded linear functionals, using the fact that the latter separate points.

Definition 3.3.3. By the open mapping theorem, the operator $T - \lambda I$ may fail to be invertible for three reasons:

(a) $T - \lambda I$ is not injective. Then, λ is an eigenvalue of T, and thus belongs to the **point spectrum** $\sigma_p(T)$.

(b) $T - \lambda I$ is injective and its range is dense in Y. Then, $(T - \lambda)^{-1}$ is an unbounded operator and λ belongs to the **continuous spectrum** $\sigma_c(T)$.

(c) If $T - \lambda I$ is injective but its range is not dense, λ is said to belong to the **residual spectrum** $\sigma_{res}(T)$.

(d) The **essential spectrum** $\sigma_{ess}(T)$ is the set of λ for which $T - \lambda I$ is not Fredholm.

Proposition 3.3.2. $\lambda \in \sigma_p(T) \iff \overline{\lambda} \in \sigma_{res}(T^*)$.

Proof.

$$\ker T - \lambda I = \operatorname{ran}(T^* - \overline{\lambda} I)^{\perp} \neq \varnothing.$$

□

Example 3.3.1. (a) Consider the left-shift operator $T : l^2 \to l^2$, i.e.

$$T((x_1, x_2, ...,)) = (x_2, x_3, ...,).$$

Note that $\|T\| = 1$. The point spectrum must satisfy

$$T(x_1, ...) = (\lambda x_1, ...,),$$

so $x_n = \lambda^{n-1} x_1$ for $n \geqslant 2$, which has a solution for any nonzero $|\lambda| < 1$. Thus, $\sigma_p(T) = \mathbb{D}$. Since $\lambda \in \sigma_p(T)$ implies $\overline{\lambda} \in \sigma_{res}(T^*)$, it follows that $\sigma_c(T) = \partial\mathbb{D}$ and $\sigma_{res}(T) = \varnothing$.

(b) Similarly, considering the right-shift operator

$$T((x_1, ...)) = (0, x_1, ...),$$

we get that $\sigma_p(T) = \varnothing$, $\sigma_{res}(T) = \mathbb{D}$, and $\sigma_c(T) = \partial\mathbb{D}$.

(c) Consider $-\Delta : H^2(\mathbb{R}^n) \subset L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$. Then, solving

$$(-\Delta - \lambda) f = g$$

equates to solving

$$(|\xi|^2 - \lambda)\widehat{f} = \widehat{g},$$

so since $f \in H^2(\mathbb{R}^n)$,

$$\lambda \in \rho(-\Delta) \iff \frac{(1 + |\xi|^2)\widehat{g}}{|\xi|^2 - \lambda} \in L^2(\mathbb{R}^n),$$

which is true whenever the multiplier is bounded, i.e. $\lambda < 0$. Thus, $\rho(-\Delta) = (-\infty, 0) \implies \sigma(-\Delta) = [0, \infty)$.

3.4 Compact and Fredholm Operators

Definition 3.4.1. TFAE:

(a) A bounded operator $T : X \to Y$ between Banach spaces sends bounded sets to relatively compact sets.

(b) If x_n is bounded sequence, Tx_n has a convergent subsequence.

In either case, $T : X \to Y$ is called a **compact operator.**

Proof. The equivalence of definitions (a) and (b) is immediate. □

Proposition 3.4.1. *A compact operator $T : X \to Y$ sends weakly convergent sequences to strongly convergent sequences. The converse holds true if X is reflexive.*

Proof. If $x_n \rightharpoonup x$ is a weakly convergent sequence, then $Tx_n \rightharpoonup Tx$ and there is some strongly convergent subsequence $Tx_{n_k} \to y = Tx$. Particularly, since every subsequence has a further subsequence converging to y, $Tx_n \to Tx = y$. Conversely, if X is reflexive, let x_n be a bounded sequence. Then, by Kakutani's theorem, x_n has a weakly convergent subsequence $x_{n_k} \rightharpoonup x$. Applying the same subsequence of a subsequence argument completes the proof. ☐

Proposition 3.4.2. *(a) A finite rank operator is compact.*

(b) $T : X \to Y$ is compact iff $T^* : Y^* \to X^*$ is compact.

(c) Compact operators form a two-sided ideal in the space of bounded operators.

(d) If $\operatorname{ran} T$ is closed in Y, T is a finite rank operator.

Proof. (a) and (c) are clear from the fact that one is dealing with bounded operators. If $\operatorname{ran} T$ is closed in Y, $T : X/\ker T \xrightarrow{\sim} \operatorname{ran} T$ is an isomorphism of Banach spaces. Since T is compact, this would contradict the fact that the unit ball in an infinite-dimensional space is not compact if $\operatorname{ran} T$ is infinite-dimensional.

Suppose T is compact, and consider $K = \overline{TB(0,1)} \subset Y$. Let $\phi_n \subset B_{Y^*}$ be a bounded sequence. Then, $\phi_n|_K$ is bounded and equicontinuous, so by Arzela-Ascoli, there is some Cauchy subsequence ϕ_{n_k}. Then, $T^*\phi_{n_k}$ is Cauchy, and so converges to some element $\psi \in X^*$. ☐

Definition 3.4.2. An operator $T : X \to Y$ is **Fredholm** if $\dim \ker T, \operatorname{codim} \operatorname{ran} T < \infty$. $\dim \ker T - \operatorname{codim} \operatorname{ran} T$ is called the **index** of T.

Lemma 3.4.1 (Riesz Lemma)**.** *If B is a vector space and $V \subset B$ is a closed proper subspace, there exists a unit vector $v \in B$ such that $d(v, V) \geqslant \alpha$ for $\alpha < 1$. If B is reflexive, then one may take $\alpha \leqslant 1$.*

If K is a compact operator, then $K(B(0,1))$ is compact, and we know that the unit ball is compact only in finite-dimensional Banach spaces. One may thus ask to what extent are compact operators different from operators that have a finite-dimensional image.

Lemma 3.4.2. *For a Hilbert spaces H, the closure of the ideal $F(H)$ of **finite-rank operators** (i.e. operators with finite-dimensional image) with respect to the norm topology in $\mathcal{B}(H)$ is the ideal $K(H)$ of compact operators.*

Proof. We will use the following characterization of compact subsets of a separable Hilbert space H : a subset $K \subset H$ is compact iff it is closed, bounded, and given an orthonormal basis $\{e_k\}$, there exists an N such that for any $u \in K$, one has $\sum_{k>N} |\langle u, e_k\rangle|^2 < \epsilon$. With this characterization in hand, one may simply define the sequence of finite rank operators $T_n u = \sum_{k\leqslant n}\langle u, e_k\rangle e_k$, and the tail condition then guarantees precisely that $T_n \to T$ in norm. Conversely, one may note that if $T_n \to T$ is a sequence of finite-rank operators, then $T_n B(0,1)$ is totally bounded, and since $T_n \to T$ in norm, $TB(0,1)$ is totally bounded, therefore precompact. ☐

Remark 3.4.1. By the same arguments, one may conclude that the ideal $K(H)$ is closed in $B(H)$ in the norm topology.

Remark 3.4.2. This is not true in general for operators $T : X \to Y$ between arbitrary Banach spaces X, Y. Spaces that satisfy the conditions of the lemma are said to satisfy the **approximation property (AP).**

Lemma 3.4.3. *If $T : X \to X$ is compact, $T - \lambda I$ is Fredholm.*

Proof. T is a multiple of the identity when restricted to $\ker(T - \lambda I)$, so it has to be finite-dimensional for its image to be compact. We show that $T - \lambda I$ is bounded below if it is injective. If not, then for some x_n such that $(T - \lambda I)x_n \to 0$, $x_n \nrightarrow 0$, $Tx_{n_k} \to y \neq 0$, one has $(T - \lambda I)Tx_{n_k} \to y = 0$, which a contradiction. Thus, $T - \lambda I$ is injective and has closed range. If $T - \lambda I$ is not injective, apply the argument on $X/\ker T$ and pull back on the image. □

Theorem 3.4.1 (Spectral Theorem for Compact Operators)**.** *If $T : X \to X$ between infinite-dimensional Banach spaces is compact, $0 \in \sigma(T) - \sigma_p(T)$, each eigenvalue has finite multiplicity, $\sigma(T)$ is countable, with the only possible limit point being the origin.*

Proof. We show $\sigma(T) = \sigma_p(T)$. If not, $T - \lambda I$ is injective but not surjective. Define $Y_n = \operatorname{ran}(T - \lambda I)^n$, and notice that since $T - \lambda I$ is injective, all these have closed range by the lemma. By the Riesz lemma, pick a sequence $y_n \in Y_n$ s.t. $d(y_{n+1}, Y_n) > 1$, and form a convergent subsequence Ty_n, note that

$$Ty_n - Ty_m = (T - \lambda I)y_n - (T - \lambda I)y_m + \lambda(y_n - y_m) \in \lambda y_n + Y_{n+1},$$

a contradiction since this implies $\|Ty_n - Ty_m\| \nrightarrow 0$. By the exact same logic, if there are infinitely many eigenvalues λ_n with eigenvectors y_n outside a ball away from the origin, if $Y_n = \operatorname{span}(y_1, ..., y_n)$, then the same logic applies, so $Ty_n - Ty_m \in \lambda_m y_m + Y_{m-1}$, i.e. $\|Ty_n - Ty_m\| > \frac{\epsilon}{2}$, a contradiction. □

Theorem 3.4.2 (Spectral Theorem for Compact Self-Adjoint Operators)**.** *If H is a (separable) Hilbert space and $T : H \to H$ is a compact, self-adjoint operator, then there exists a (countable) orthonormal basis of eigenvectors with real eigenvalues for H, i.e. T is unitarily diagonalizable.*

Remark 3.4.3. The separability of H is needed for the basis of $\ker T$ to be countable.

Proposition 3.4.3. *$T : X \to Y$ is Fredholm iff there exists $S : Y \to X$ such that $I - TS, I - ST$ are compact.*

Proof. If T is Fredholm, let S be the composition of the projection from Y onto $\operatorname{ran} T$ and the isomorphism from $\ker T^{\perp}$ and $\operatorname{ran} T$. Then, $I - ST, I - TS$ are easily verified to be finite rank projections, hence compact. The converse follows since $T - \lambda I$ for compact T is Fredholm. □

Corollary 3.4.1. *If T is Fredholm and K is compact, $T + K$ is Fredholm.*

Proposition 3.4.4 (Weyl)**.** *Let $T : X \to X$ be self-adjoint and $K : X \to X$ be self-adjoint compact. Then, $\sigma_{ess}(T) = \sigma_{ess}(T + K)$.*

Proof. This follows from the fact that being Fredholm is invariant under relatively compact perturbations. □

3.4.1 Exercises

Problem 3.4.1 (Spring 2010 Problem 13). Suppose X, Y are Banach spaces, and X is separable and X^* is separable. Show that $T : X \to Y$ is compact iff for every bounded sequence $x_n \in X$, there exists a subsequence x_{n_k} and a $\phi \in X$ such that $x_{n_k} = \phi + r_{n_k}$, where $Tr_{n_k} = 0$.

Proof. For the backward direction, note that $Tx_{n_k} \to T\phi$, so T is compact. Conversely, note that since X^* is separable, the unit ball in X^{**} is metrizable with respect to the weak-*

topology, i.e. weak-* compactness and weak-* sequential compactness are equivalent. Then, by Banach-Alaouglu, X^{**} is weak-* compact and therefore weak-* sequentially compact, i.e. for any bounded sequence $x_n \in X$, there is a subsequence $\widehat{x_{n_k}} \xrightarrow{*} \hat{x}$ in X^{**}, where $\hat{x} = J(x)$ for some $x \in X$ since X is reflexive. Then, $\phi(x_{n_k}) \to \phi(x)$ for all $\phi \in X^*$, i.e. $x_{n_k} \rightharpoonup x$. Since compact operators map weakly convergent sequences to strongly convergent subsequences, we set $\phi = x$ and obtain $T(x_{n_k} - \phi) \to 0$ in Y. □

Problem 3.4.2 (Problem 6 Fall 2014). Let X be a Banach space. Show that if X^* is separable, then X is separable.

Proof. Let $\{\phi_n\}$ be a dense sequence in the unit sphere of X^*. Then for any $q \in \mathbb{Q}$, there exists an $x_{q,n}$ in the unit ball of X^* such that $\phi_n(x_{q,n}) = q$. Let $S = \text{span}_{\mathbb{Q}}(x_{q,n})$, which is clearly countable. I claim that S is dense in the unit ball of X. Indeed, S is clearly dense in $\text{span}_{\mathbb{R}}(x_{q,n})$, so it suffices to show the latter is dense in the unit ball. Suppose it is not. Then, by Hahn-Banach, there is a linear functional ϕ in the unit sphere of X^* that vanishes on S yet does not vanish for some x on the unit sphere. Then, for ϕ_n such that $\|\phi - \phi_n\| < \frac{1}{3}$, $|\phi(x_{\frac{1}{2},n}) - \phi_n(x_{\frac{1}{2},n})| = \frac{1}{2} > \frac{1}{3}$, which is a contradiction. □

Problem 3.4.3 (Problem 5 Spring 2014). Prove that l^1 and l^2 are separable but l^∞ is not. Moreover, show there is no bounded surjective map from l^2 to l^1.

Proof. The $\mathbb{Q}$-span of unit vectors in l^1 and l^2 is separable, for any $a \in l^1$, one may pick $x = (x_1, ...x_n, 0, ...)$ such that $\|(a_n)_n^\infty\|_p < \epsilon$ and $\sum_{k=1}^n |x_k - a_k|^p < \epsilon$. However, l^∞ is not separable. Suppose it was. Then the unit ball in l^∞ is separable. It suffices to show that one can find uncountably many elements that are all at least one away from each other. But clearly, if we take the subset of distinct binary sequences, it is uncountable and each element is a distance of exactly one away from all others. Thus, l^∞ is not separable.
Now, suppose there is a bounded surjective map T from l^2 to l^1. Then, the adjoint T^* is a bounded map $T^* : l^\infty \to l^2$. Moreover, $\ker T^{*\perp} = \overline{\text{im}(T)}$, i.e. $T*$ is injective. However, l^2 is separable while l^∞ is not, contradicting the existence of an injective continuous map $l^\infty \to l^2$. Thus, there is no bounded surjective map from l^2 to l^1. □

Problem 3.4.4 (Problem 6 Fall 2014). Let a_n be a sequence of elements in a Hilbert space H such that $\|a_n\| = 1$ for all n. Show that if the span of $\{a_n\}$ spans H, then H is finite-dimensional. Moreover, show that if $a_n \rightharpoonup 0$, then 0 is in the closed convex hull of $\{a_n\}$.

Proof. Note that if H was infinite-dimensional, it would have countable dimension as a vector space. But a standard Baire category theorem argument yields that any Banach space has uncountable dimension as a vector space. Thus, H is finite dimensional.
Now, suppose that $a_n \rightharpoonup 0$. We want to show that for some $t_{n_1} + ... + t_{n_k} = 1$, $\sum_{i=1}^k t_{n_i} a_{n_i} = 0$. □

Problem 3.4.5. Let $L_n \in L^{\infty *}$ be a sequence of functionals defined by $L_n(f) = \frac{1}{n!}\int_0^\infty x^n e^{-x} f(x)dx$. Show that $\{L_n\}$ has no weak-* convergent subsequence. Why does this not contradict Banach-Alaouglu?

Proof. This does not contradict Banach Alaoglu as the compactness of the weak-* topology on the unit ball in $L^{\infty *}$ implies sequential compactness only when the topology on the unit ball is metrizable, which is not necessarily the case (since L^∞ is not separable).
We now show that there is no weak-* convergent subsequence by constructing a function f such that $L_n(f) \sim (-1)^n$. For sake of contradiction, suppose there is a convergent subsequence. Let $I_{n_k} = [a_{n_k}, b_{n_k}]$ be a sequence of disjoint intervals such that $b_{n_k} < a_{n_{k+1}}$, where n_k is a subsequence of the chose subsequence, such that $L_{n_k}(\chi_{I_{n_j}}) \geqslant 1 - \epsilon$ and $L_{n_k}(\chi_A) < \epsilon$ for some $\epsilon > 0$ for any A such that $A \cap I_{n_j} = \varnothing$. This is possible since $f_n = \frac{x^n e^{-x}}{n!}$ converges to 0 pointwise, and it is monotonically increasing on $[0, n]$ and monotonically decreasing on $[n, \infty)$, so on any interval $[a, b]$, the sequence f_n is bounded by $\sup(a, b)\chi_{[a,b]}$, i.e. by dominated convergence theorem, $L_n(\chi_{[a,b]}) \to 0$ as $n \to \infty$. Particularly, since $\|L_n\| = 1$, it implies that the mass of f_n is concentrated further away from the origin as $n \to \infty$, so it is possible to find such a sequence of intervals by taking an appropriate subsequence n_k. Then, $f = \sum_{k=1}^\infty (-1)^k \chi_{I_{n_k}} \in L^\infty$ has norm 1 in L^∞, so

$$|L_{n_k}(f) - (-1)^{n_k}| \leqslant |\epsilon + L_{n_k}(f - \chi_{I_{n_k}})| \leqslant 2\epsilon,$$

so we get a contradiction. □

3.5 Banach Algebras

We are now interested in introducing an algebra structure on Banach spaces:

Definition 3.5.1. A **(real/complex) Banach algebra** A is a Banach space that is given the structure of an algebra over $\mathbb{R}$ or $\mathbb{C}$, where the multiplication map $(x, y) \to xy$ is continuous, or (which can be seen by uniform boundedness and rescaling to an equivalent norm) equivalently, satisfying $\|xy\| \leqslant \|x\|\|y\|$ for all $x, y \in A$.

Remark 3.5.1. Without loss of generality, one may assume the algebra is unital, as for a nonunital algebra A, the algebra $A \times \mathbb{K}$ with multiplication given by $(a, z_1)(b, z_2) = (ab + az_2 + bz_1, z_1 z_2)$ is Banach algebra with unit $(0, 1)$, with $A \hookrightarrow A \times \mathbb{K}$ being an isometric embedding.

Example 3.5.1. The space $\mathcal{B}(X)$ of bounded operators $T : X \to X$ on a Banach space is the prototypical example of a unital Banach algebra.

With the additional structure of an algebra, the spectrum of A now has additional useful properties.

Definition 3.5.2. A **character** is an algebra homomorphism $\phi : A \to \mathbb{K}$. The space of characters of A is denoted as $\Delta(A)$.

Proposition 3.5.1. $\Delta(A)$ *is a weak-* compact subset of the unit sphere in* A^*.

Proof. Note that b being invertible implies $\alpha(b)$ is invertible with inverse $\alpha(b^{-1})$. If $\|\alpha\| \neq 1$, there exists b s.t. $\|b - 1\| < 1$ but $\alpha(b - 1) = \alpha(b) - 1 > 1$ (or vice versa). Then,

$$\alpha\left(\frac{b - 1}{\alpha(b) - 1} - 1\right) = 0,$$

but the element inside is invertible since the ratio (up to flipping the fraction) has norm less than 1, which is a contradiction. It is easy to check that it is weak−* closed and therefore weak-* compact. □

Proposition 3.5.2. *Every maximal ideal of A is closed. Moreover, if A is a commutative, then there is a bijection between the maximal ideals of A and $\Delta(A)$.*

Proof. Recall that the set of invertible elements is open (by power series expansion). If $\mathfrak{m}$ is a maximal ideal of A, its closure is easily shown to be a maximal ideal containing $\mathfrak{m}$, so by maximality it equals $\mathfrak{m}$.

We briefly prove an important lemma:

Lemma 3.5.1 (Gelfand-Mazur Theorem)**.** *If a normed algebra A (over $\mathbb{R}/\mathbb{C}$) is a division algebra, then $A = \mathbb{R}, \mathbb{C}, \mathbb{H}$ (or $A = \mathbb{C}$).*

Proof. The general case is complicated, so for simplicity, we suppose that A is a complex Banach algebra. A classic result is that the spectrum of an element in a complex Banach algebra is nonempty. But $a - \lambda I$ is noninvertible iff $a - \lambda I = 0$, i.e. $a = \lambda I$. □

Now, if A is commutative and I is a maximal ideal, then A/I is a field over $\mathbb{C}$, so by the Gelfand-Mazur theorem, $A/I = \mathbb{C}$. Finally, we note that by the first isomorphism theorem, if $\alpha \in \Delta(A)$, $\ker \alpha$ is a closed maximal ideal of codimension 1. Consequently, for commutative Banach algebras, we have the following bijection:

$$\begin{aligned} \Delta(A) &\Longleftrightarrow \text{MaxSpec}(A) \\ \alpha &\to \ker \alpha \\ \alpha : (a \to a\mathfrak{m}) &\leftarrow \mathfrak{m} \end{aligned}$$

□

Recall that for any Banach space E, we have an embedding $E \hookrightarrow C(X)$, where X is the unit ball of E^* and is therefore weak-* compact. Motivated by this, we consider the map $A \to C(\Delta(A))$. In general, this may not be injective, surjective, or isometric. However, it turns out that if we restrict ourselves to a special class of Banach algebras, this map turns out to be an isometric isomorphism.

Definition 3.5.3. A **$*-$algebra** is an algebra equipped with an anticommutative involution $*$ satisfying $(xy)^* = y^*x^*$. If A is a Banach $*-$algebra satisfying $\|x^*x\| = \|x\|^2$ for all $x \in A$, then A is called a **C^*-algebra.** A homomorphism preserving the involution is called a **$*$-homomorphism.**

Remark 3.5.2. Since $\|xy\| \leqslant \|x\|\|y\|$, the last condition is equivalent to $\|xx^*\| = \|x\|\|x^*\|$ for all $x \in A$.

Example 3.5.2. The prototypical example of a C^*-algebra is that of $\mathcal{B}(H)$, the space of bounded linear operators on a Hilbert space H.

The map $\Phi : A \to C(\Delta(A))$ defined earlier is known as the **Gelfand transform.**

Definition 3.5.4. The **spectral radius** of a is $r_a = \sup_{z \in \sigma(a)} |z|$.

Lemma 3.5.2 (The spectral radius formula)**.** $r_a = \lim_{n\to\infty} \|a^n\|^{\frac{1}{n}}$.

Proof. Note that the the power series for $(a-\lambda I)^{-1} = \lambda^{-1}(\frac{a}{\lambda}-I)^{-1}$ converges whenever $|\lambda^{-1}| < \|a\|^{-1}$ and diverges whenever $|\lambda^{-1}| > \limsup_{n\to\infty} \|a^n\|^{-\frac{1}{n}}$. This means that the spectral radius is at least $r_a \geqslant \limsup_{n\to\infty} \|a^n\|^{\frac{1}{n}}$. Conversely, by factoring $a^n - \lambda^n$, one sees that $\lambda \in \sigma(a) \implies \lambda^n \in \sigma(a^n)$, so $r_a^n \leqslant \|a^n\|$, and by taking the liminf on both sides, the claim follows. □

Corollary 3.5.1. *If a is normal in a C^*-algebra, then*

$$r_a^2 \leqslant \|a\|^2 = \|a^*a\| \leqslant \lim_{n\to\infty} \|a^{*n}\|^{\frac{1}{n}} \|a^n\|^{\frac{1}{n}} = r_a^2,$$

i.e. $r_a = \|a\|$.

Corollary 3.5.2. *Since $\|a\|^2 = \|a^*a\| = r_{a^*a}$, there is a unique norm on a *-algebra that makes it into a C^*-algebra.*

Lemma 3.5.3. *If a is self-adjoint, $\sigma(a) \subset \mathbb{R}$.*

Proof. One easily verifies that if a is self-adjoint, then e^{ia} is unitary. Then, for $\lambda \in \sigma(a)$,

$$e^{ia} - e^{i\lambda} = (e^{i(a-\lambda)} - 1)e^{i\lambda} = (a-\lambda)be^{i\lambda}$$

for $b = \sum_{n=1}^{\infty} \frac{i^n(a-\lambda)^{n-1}}{n!}$. b commutes with a and $a - \lambda$ is not invertible, so $e^{ia} - e^{i\lambda}$ is not invertible. If u is unitary, $\|u\|^2 = \|u^*u\| = 1$, i.e. $r_u = 1$, and since $\lambda \in \sigma(u) \implies \lambda^{-1} \in \sigma(u^{-1})$, we have that $\sigma(u) \subset S^1$. Since e^{ia} is unitary, we thus have $|e^{i\lambda}| = 1$, i.e. $\lambda \in \mathbb{R}$. □

We are now ready to state the main results of this section.

Theorem 3.5.1 (Gelfand Representation Theorem)**.** *If A is a commutative C^*-algebra, the Gelfand transform is an isometric *-isomorphism.*

Proof. Recall that $\Phi(a)(\alpha) = \alpha(a)$. By the identity

$$a = \frac{a + a^*}{2} + i\frac{(-ia) + (-ia)^*}{2},$$

it suffices to show Φ is a *-homomorphism on self-adjoint elements, i.e. $\phi(a^*) = \overline{\phi(a)}$ for $\phi \in \Delta(A), a \in A$ self-adjoint. But this follows from the fact that $\sigma(a) = \sigma(a^*) = \operatorname{Im}\Phi(a) \subset \mathbb{R}$.

By the C^* identity, we have that $\|a\| = \|a^{2n}\|^{\frac{1}{2^n}} \to r_a = \|\Phi(a)\|_\infty$, so Φ is an isometry and is therefore injective. Note that $\operatorname{Im}\Phi$ clearly separates points, since if $\alpha_1(a) - \alpha_2(a) = 0$ for all a, then $\alpha_1 = \alpha_2$. Then, by Stone-Weierstrass and the fact that the image of an isometry is closed, Φ is surjective and therefore an isometric *-isomorphism. □

Corollary 3.5.3. *This shows that every $\phi \in \Delta(A)$ is in fact a *-homomorphism.*

Corollary 3.5.4. *If $a \in A$ is normal, then $\Phi : C^*(\{a\}) \to C(\sigma(a)\backslash\{0\})$ is an isometric *-isomorphism s.t. $\Phi(a) = id_{\sigma(a)\backslash\{0\}}$. This is because $\phi \in \Delta(A)$ is uniquely determined by the image of $0 \neq \lambda = \phi(a) \in \sigma(a)$ (by the *-homomorphism property of $\Delta(A)$), i.e. $\Delta(A) \cong \sigma(a)\backslash\{0\}$, and $\Phi(a)(\phi) = \Phi(a)(\phi(a)) = \phi(a)$, showing that $\Phi(a)$ is the identity map.*

Definition 3.5.5 (Continuous Functional Calculus)**.** For $f \in C(\sigma(a))$, define $f(a)$ to be the element corresponding to f in the above isomorphism. Note that by the $^*-$homomorphism property of $\Delta(A)$, $f(a) = p(a)$ for all polynomials $p(z)$. Additionally, since Φ is an isometry, $f_n \to f$ uniformly implies that $f_n(a) \to f(a)$.

Corollary 3.5.5 (Spectral Mapping Theorem)**.** *If a is a normal element in a C^*-algebra and $f \in C(\sigma(a))$, the corollary above implies that* $\operatorname{Im}\Phi(f(a)) = f(\sigma(a))$, *i.e.* $\sigma(f(a)) = f(\sigma(a))$.

Proof. Note that

$$\lambda \in \sigma(a) \iff a - \lambda I \text{ not invertible} \iff \Phi(a - \lambda I) \text{ not invertible} \iff \lambda \in \operatorname{Im}\Phi(a).$$

This implies that $\sigma(a) = \operatorname{Im}\Phi(a)$, and the rest follows immediately from the Gelfand representation theorem. □

We now list some applications of the continuous functional calculus.

Proposition 3.5.3. *TFAE:*

(a) a is self-adjoint and $\sigma(a) \subset [0, \infty)$.

(b) $a = b^2$ for some self-adjoint b.

(c) a is self-adjoint and $\|\lambda - a\| \leqslant |\lambda|$ for some/any $\lambda \geqslant \|a\|$.

Proof. The forward direction is immediate by applying the spectral mapping theorem to $f(x) = \sqrt{x}$ and the fact that the continuous functional calculus commutes with the * operation. The backward direction is immediate by applying the spectral mapping theorem to $f(x) = x^2$. Finally,

$$\|\lambda - a\| = \sup_{z \in \sigma(a)} |\lambda - z| \leqslant \lambda$$

for $\sigma(a) \subset [0, \infty)$, and the converse can be established similarly. □

Definition 3.5.6. If any of these hold, a is called a **positive element,** denoted $a \geqslant 0$. We denote A_+ as the **cone (i.e. a vector space closed under positive scalars)** of positive elements. Note that this is indeed a cone by part (c) of the above characterization and the triangle inequality.

Example 3.5.3. If $|a| := \sqrt{a^2}$, $a_\pm := \frac{1}{2}(|a| \pm a)$, then $a_\pm, |a| \geqslant 0$ and $a_+a_- = a_-a_+ = 0$.

Lemma 3.5.4. *If* $0 \leqslant a \leqslant b, \|a\| \leqslant \|b\|$. *If a, b are also invertible, then* $0 \leqslant b^{-1} \leqslant a^{-1}$.

Proof. Note in general that if a, b commute, and f_a, f_b are the continuous functions corresponding to a, b, $a \leqslant b \iff f_a \leqslant f_b$. Then, $f_b \leqslant f_{\|b\|}$ as functions on $C^*(\{b\})$, so $a \leqslant b \leqslant \|b\|$, i.e. $f_a \leqslant f_{\|b\|}$, so $\|a\| = \|f_a\|_\infty \leqslant \|b\|$. For the second property, note that $a \leqslant b \implies c^*ac \leqslant c^*bc$, since $A_+ = \{a^*a : a \in A\}$. Thus, $a \leqslant b$ implies $b^{-\frac{1}{2}}ab^{-\frac{1}{2}} \leqslant 1$, so

$$\|a^{\frac{1}{2}}b^{-1}a^{\frac{1}{2}}\| = \|b^{-\frac{1}{2}}a^{\frac{1}{2}}\|^2 = \|b^{-\frac{1}{2}}ab^{-\frac{1}{2}}\| \leqslant 1,$$

i.e. $a^{\frac{1}{2}}b^{-1}a^{\frac{1}{2}} \leqslant 1$. Thus,

$$b^{-1} = a^{-\frac{1}{2}}(a^{\frac{1}{2}}b^{-1}a^{\frac{1}{2}})a^{-\frac{1}{2}} \leqslant a^{-1}.$$

□

The main theorem of C^*-algebra states that the canonical example of a C^*-algebra is in fact the only possible case, up to isomorphism.

Theorem 3.5.2 (Gelfand-Naimark-Segal (GNS) Theorem)**.** *Every C^*-algebra is isometrically *-isomorphic to a closed subalgebra of $\mathcal{B}(H)$ for some Hilbert space H.*

Another important class of algebras are the so-called von Neumann algebras.

Definition 3.5.7. A **von Neumann algebra (or a W^*-algebra)** is a C^*-algbra A s.t. there exists an algebra B s.t. $B^* = A$, i.e. A has a **predual.** Equivalently, A is a weakly closed subalgera of $\mathcal{B}(H)$.

Theorem 3.5.3 (Von Neumann Bicommutant Theorem)**.** *If $A \subset \mathcal{B}(H) := G$ is a *-subalgebra, then the von Neumann algebra generated by A, i.e. the weak closure of A, is equal to the strong closure of A and is equal to $A'' = C_G(C_G(A))$, the bicommutant of A.*

Proof. Note that $A \subset A''$. If $T \notin A'$, there exists an operator S and a weak neighborhood where $\langle TSx, y\rangle - \langle STx, y\rangle \neq 0$, i.e. A' is weakly closed. Thus, the weak closure of A is contained in A''. Since strongly closed implies weakly closed, it follows that the strong closure is a subset of the weak closure. Finally, it suffices to show that any open neighborhood of $T \in A''$ in the strong topology contains an element of A. Note that for $h \in H$, the norm closure H_1 of Ah is a closed subspace of G, and let P be the corresponding projection.

Lemma 3.5.5. *$P \in A'$.*

Proof. For $x \in H$, let $O_n h \to Px$. For $S \in A$, we thus have $SO_n h \to SPx \in H_1$, so $PSPx = SPx$, i.e. $PSP = SP$ for all $S \in A$. Finally, using the fact that A is a *-subalgebra, we get that

$$\langle x, SPy\rangle = \langle x, PSPy\rangle = \langle PS^*Px, y\rangle = \langle x, PSy\rangle,$$

so $PS = SP$, i.e. $P \in A'$. ☐

Then, $Th = TPh = PTh \in H_1$, so by the definition of H_1, $\|Th - Sh\| < \epsilon$ for some $S \in A$ for any $\epsilon > 0$. Thus, there exists a sequence $S_n \to T$ in the strong topology, so T is in the strong closure of A. ☐

3.6 Borel and Holomorphic Functional Calculus

We have seen that for a C^*-algebra, we can define a continuous functional calculus. We are interested in defining a more restrictive functional calculus on a more general class, namely, just Banach algebras.

Definition 3.6.1. A function $f : U \subset \mathbb{C} \to B$ with values in a Banach space B is **analytic/holomorphic** if $\lim_{z\to z_0} \frac{f(z)-f(z_0)}{z-z_0}$ exists for all $z \in U$. f is **weakly analytic** if $\phi \circ f$ is analytic for all $\phi \in B^*$.

Remark 3.6.1. If f is continuous, weakly analytic $\iff$ analytic.

Definition 3.6.2. For a function $f : X \to B$ with values in a Banach space B, one may define the **Bochner integral** the same way one defines the Lebesgue integral on $\mathbb{R}$. In particular, if $s_n \to f$ is a sequence of simple functions, one may define $\int f = \lim \int s_n$, which can be shown to be Cauchy and independent of the chosen sequence. The Bochner integral satisfies $\int Tf = T\int f$ for any bounded linear operator T, and one has all the classic complex and real analytic results (DCT, Monotone Convergence, Fatou, Cauchy's Theorem/Integral Formula).

Remark 3.6.2. Radon-Nikodym property.

Definition 3.6.3 (Holomorphic Functional Calculus)**.** For $f : U \to \mathbb{C}$ holomorphic, $a \in B$, where B is a Banach algebra and $a \in B$, and $\sigma(a) \subset U$, define

$$f(a) := \frac{1}{2\pi i}\int_\gamma \frac{f(z)}{z-a}dz,$$

where $\gamma = B(0, r_a + 1)$ encloses $\sigma(a)$ and $\frac{1}{z-a}$ is the resolvent mapping.

Proposition 3.6.1. *(a) The integral is well-defined and independent of the choice of γ.*

(b) $(fg)(a) = f(a)g(a)$, i.e. $f \to f(a)$ is a homomorphism.

(c) If $f_n \to f$ normally, then $f_n(a) \to f(a)$.

(d) $id(a) = a$ and the calculus agrees with the continuous functional calculus on a C^*-algebra.

Proof. The well-definedness is a consequence of Cauchy's integral formula for the Bochner integral. The fact that the map is a homomorphism follows from Fubini. Indeed, if we assume that γ_1 is in the interior of the region bounded by γ_2,

$$\begin{aligned} f(a)g(a) &= \frac{1}{(2\pi i)^2}\int_{\gamma_1}\int_{\gamma_2} f(z_1)g(z_2)\frac{(z_1-a)^{-1}-(z_2-a)^{-1}}{z_2-z_1}dz_1dz_2 \\ &= \frac{1}{(2\pi i)^2}\left[\int_{\gamma_1}\frac{f(z_1)}{z_1-a}\left[\int_{\gamma_2}\frac{g(z_2)}{z_2-z_1}dz_2\right]dz_1 - \int_{\gamma_2}\frac{g(z_2)}{z_2-a}\left[\int_{\gamma_1}\frac{f(z_1)}{z_2-z_1}dz_2\right]dz_1\right] \\ &= \frac{1}{2\pi i}\int_{\gamma_1}\frac{f(z_1)}{z_1-a}\left[\frac{1}{2\pi i}\int_{\gamma_2}\frac{g(z_2)}{z_2-z_1}dz_2\right]dz_1 = \frac{1}{2\pi i}\int_{\gamma_1}\frac{f(z_1)g(z_1)}{z_1-a}dz_1 = (fg)(a), \end{aligned}$$

where the second term vanishes since $z_1\gamma_2$ is outside the region bounded by γ_1. Normal convergence implies convergence in norm immediately by definition, and

$$\frac{1}{2\pi}\int_\gamma \frac{z}{z-a}dz = \sum_{n\geqslant 0} a^n \int_\gamma \frac{1}{z^n}dz = a$$

by basic properties of complex integrals. Finally, for continuous f, let $p_n \to f$ be a normally convergent sequence of polynomials. Then, since the holomorphic and continuous calculi agree on polynomials, by the uniform convergence property, $p_n(a) \to f(a)$, which is the same element in both functional calculi. □

Remark 3.6.3. Note that the proof requires that f be holomorphic on $\sigma(a)$ as well, so one cannot apply the calculus to, for instance, meromorphic functions with poles in $\sigma(a)$.

Notice that the curve γ encloses the entire spectrum. If it did not enclose any of the spectrum, the resolvent would be analytic and the integral would evaluate to 0 by Cauchy's theorem. But what if we include only part of the spectrum?

Definition 3.6.4. For a curve γ avoiding $\sigma(a)$ and enclosing some compact subset $K \subset \sigma(a)$, define the **Riesz projector**

$$\Pi_K := \frac{1}{2\pi i}\int_\gamma \frac{1}{z-a}dz.$$

Proposition 3.6.2. *(a) Π_K is a projection that commutes with a, and if a is self-adjoint, Π_K is an orthogonal projection (i.e. self-adjoint and s.t. $\Pi_K^2 = \Pi_K$).*

(b) If a is a bounded operator on a Banach space, $\operatorname{Im}\Pi_K, \operatorname{Im}(I-\Pi_K)$ are disjoint a-invariant subspaces with $\sigma(a|_{\operatorname{Im}\Pi_K}) = K, \sigma(a|_{\operatorname{Im}(I-\Pi_K)}) = \sigma(a)\backslash K$.

Proof. (a) Following the calculation in Proposition 5.12 with $f \equiv g \equiv 1$ and $\widetilde{\gamma}$ being a perturbation containing γ, we obtain that $g(z_2) = 1$, so $\Pi_K^2 = \Pi_K$. If a is self-adjoint, then $((z-a)^{-1})^* = (\overline{z}-a)^{-1}$, so letting γ be a circular contour and reparametrizing to go in the counterclockwise direction yields $\Pi_K^* = \Pi_K$. The fact that Π_K commutes with a is analogous to proof of the homomorphism property of the holomorphic functional calculus.

(b) The commutativity implies that $\operatorname{Im}\Pi_K, \operatorname{Im}(I-\Pi_K)$ are a-invariant, and it is easy to check that they are disjoint. Note that $\rho(A) \subset \rho(A\Pi_K)$, since the resolvent of A restricted to $\operatorname{Im}\Pi_K$ is the resolvent of $A\Pi_K$. Applying the same claim to $I-\Pi_K$ yields $\rho(A\Pi_K) \subset \rho(A)$.

□

Remark 3.6.4. This suggests that the Riesz projector is to be interpreted as a projection onto the corresponding part of $\sigma(a)$.

Of particular interest is then the case whenever a is a compact operator, since the spectral theorem yields a discrete spectrum in this case.

Theorem 3.6.1. *If $T : X \to X$ is a self-adjoint operator with discrete spectrum (e.g. when T is compact), then one may write $X = \bigoplus \operatorname{Im}\Pi_{\lambda_i}$, where Π_{λ_i} are the projections onto* $\ker(T-\lambda_i)$, *yielding the spectral decomposition*

$$X = \bigoplus_i \ker(T-\lambda_i).$$

Corollary 3.6.1. *This immediately yields the decomposition into invariant subspaces for finite-dimensional spaces, which is known as the Jordan canonical form.*

Now that we have developed two functional calculi, we are ready to formalize the most general functional calculus for Banach algebras, known as the Borel functional calculus. For this, we need the most general version of the spectral theorem.

Theorem 3.6.2 (Spectral Theorem, General Version)**.** *The Gelfand embedding $\Phi : C(\sigma(A)) \hookrightarrow \mathcal{B}(H)$ for a C^*-algebra $A \subset \mathcal{B}(H)$ may be extended to a $^*-$homomorphism $L^\infty(\sigma(A)) \to \mathcal{B}(H)$.*

Proof. We sketch the proof. Define $\psi_{x,y}(f) = \langle \Phi(f)x, y\rangle = \int f d\mu_{x,y}$ for $f \in C(\sigma(A))$, where $\mu_{x,y}$ is the measure given by the Riesz-Markov representation theorem. Moreover, $f \to \int f d\mu_{x,y}$ is bounded and bilinear for $f \in L^\infty$. Thus, there exists a unique A_f s.t. $\int f d\mu_{x,y} = \langle A_f x, y\rangle$. One may then arduously verify that $f \to A_f$ is a n injective *-homomorphism. □

Remark 3.6.5. Given $T \in \mathcal{B}(H)$ normal, the spectral theorem corresponding to the Gelfand embedding of $C^*(\{T, T^*\})$ yields the the maps $\nu_\xi : E \to \langle A_{\chi_E}\xi, \xi\rangle$, which are collection of measures called the **spectral measures of T associated to ξ**. For $\|\xi\| = 1$, these are probability measures.

Definition 3.6.5 (Borel Functional Calculus)**.** The embedding $L^\infty(\sigma(A)) \to \mathcal{B}(H)$ is called the **Borel functional calculus,** defining $f(T) \in \mathcal{B}(H)$ for a normal operator T.

Example 3.6.1. Recall that $\Delta : H^2 \to L^2$. Consider the Schrodinger equation

$$iu_t = -\Delta u, u(0) = u_0 \in H^2.$$

We may formally define the solution

$$u(t) = e^{it\Delta}u_0,$$

where we may now rigorously define the unbounded densely-defined operator $e^{it\Delta} : L^2 \to L^2$ as

$$e^{it\Delta}u = \int_\gamma \frac{e^{itz}}{(z-\Delta)^{-1}} dz,$$

which may be equivalently expressed in terms of the Fourier transform as

$$(e^{it\Delta}u)(x) = \mathcal{F}^{-1}(e^{it\xi^2}\widehat{u}) = \frac{1}{(4\pi it)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{i\frac{|x-y|^2}{4t}} u(y)dy.$$

By dominated convergence, one may see that for $u_0 \in L^1 \cap L^2$, u is continuous and satisfies the **dispersive estimate**

$$\|u(t)\|_\infty \leqslant \frac{\|u_0\|_1}{|4\pi t|^{\frac{n}{2}}}.$$

3.7 Baire Category Theorem

Recall the Baire Category Theorem:

Definition 3.7.1. A subset X of a topological space is **meager or of first category** if it is a countable union of nowhere dense sets. Otherwise, X is called **nonmeager or second category.** The complement of a meager set is called **comeager or residual.**

Remark 3.7.1. The closure of a nowhere dense set is nowhere dense.

Theorem 3.7.1 (Baire Category Theorem)**.** *TFAE:*

(a) A topological space X is nonmeager.

(b) A comeager set is dense.

(c) A countable intersection of open dense sets is dense.

*If any of these conditions hold in a topological space X, X is called a **Baire space.** Moreover, every complete metric space and locally compact Hausdorff space is a Baire space.*

Proof. We prove the equivalences. Note that the complement of an open dense set is a nowhere dense set. Then, the last statement implies that the union of nowhere dense sets cannot be the entire space. A countable intersection of open dense sets is a comeager set. Conversely, if A is comeager, $\overline{A^c}^c \subset A$ is the countable intersection of open dense sets, and is therefore dense.

Now, suppose X is a complete metric space, U_i a collection of open dense sets, and let $x_1 \in A \subset X$ be an open set. Then $A \cap U_1$ contains a closed ball. Proceeding inductively, we choose a sequence of balls $B(x_i, \epsilon_i) \subset B(x_{i-1}, \epsilon_{i-1}) \cap U_{i-1}$ for $\epsilon_i \to 0$. Then, the sequence is Cauchy, converging to $x \in A \cap \bigcap_i U_i$, so $\bigcap_i U_i$ is dense in X. ☐

Remark 3.7.2. It is very important to note that even though meager sets are countable unions of nowhere dense sets, meager sets can still be dense. For instance, $\mathbb{Q} \subset \mathbb{R}$ is meager and dense, and so is $C^1([0,1]) \subset C([0,1])$. Meanwhile, comeager sets are always dense (in a complete metric space).

Example 3.7.1. (a) $\mathbb{Q}$ is an F_σ but not G_δ set, for if $\mathbb{Q} = \bigcap_i U_i$, for some enumeration q_i of the rationals, $\bigcap U_i \backslash \{q_i\} = \varnothing$, contradicting Baire category theorem. Similarly, $\mathbb{R}\backslash\mathbb{Q}$ is G_δ but not F_σ.

(b) A Banach space must have uncountable dimension as a vector space. Otherwise, it is the union of finite-dimensional subspaces, which are nowhere dense, which contradicts the Baire category theorem.

(c) Consider the subspace $X \subset C([0,1])$ of functions differentiable at some point x. Define

$$A_{n,m} = \{f \in X : \exists x, |x - t| < \frac{1}{m} \implies \left|\frac{f(x) - f(t)}{x - t}\right| \leqslant n\}.$$

By Bolzano-Weierstass, one can verify that $A_{n,m}$ is closed. One can also show that $A_{n,m}$ has empty interior and so is nowhere dense, so by Baire category, we conclude that $X \subset \bigcup A_{n,m}$ is meager in $C([0,1])$. Thus, "most" continuous functions are in fact nowhere differentiable.

(d) (*) If $f \in C^{\infty}([0,1])$ is such that $f^{(n)}(x) = 0$ for each x for large enough n dependent on x, then f is a polynomial. Indeed, by way of contradiction, consider the sets

$$X_n = \{f^{(n)}(x) = 0\}, S = \{x : f \text{ not a polynomial on any open interval containing } x\}.$$

Note that S is nonempty and closed. Applying the Baire category theorem to X, we get that $(a,b) \cap S \subset X_n \cap S$ for some n. at least one of them must have nonempty interior. Now, on any open subset of $(a,b)\backslash S$, by the definition of S and X_n, f has to be some polynomial of degree $d < n$. But then $(a,b)\backslash S \subset X_n$, so f is a polynomial of degree at most n on (a,b), a contradiction.

The real power of Baire category theorem is to provide an elementary proof of the Open Mapping Theorem from functional analysis.

Theorem 3.7.2. *If $T : X \to Y$ is a bounded linear map between Banach spaces, either T is surjective and open, or its image is of first category in Y.*

Proof. If T is surjective, the we note that by the Baire category theorem $TB(0,n)$ has nonempty interior for some $n \in \mathbb{N}$. In particular, it must contain an open neighborhood of the origin, which completes the proof. □

Remark 3.7.3. Thus, the approach is as follows. If $V \subset W$ is a closed proper subspace, as long as the inclusion map is continuous, the above theorem implies that V is of first category in W. If V is open, one must describe V as a union of closed subspaces, each of which is meager.

3.8 Borel Sets

Remark 3.8.1. Recall that the **product topology** (as opposed to the **box topology**) is the topology generated by rectangles with finitely many nontrivial components (and is more useful since it satisfies **Tychonoff's theorem,** which says that any product of compact topological spaces is compact).

Definition 3.8.1. The **product σ-algebra** on a product of measurable topological spaces is the smallest σ-algebra that makes the projection maps π_i, measurable, i.e. the σ-algebra generated by **cylinder sets** - products of at most finitely many nontrivial **measurable** sets. If the product is countable, then the product σ-algebra is in fact generated by arbitrary **rectangles**, that is, products of elements of the respective σ-algebras.

Proposition 3.8.1. *If $\mathcal{B}(\mathbb{R}), \mathcal{L}(\mathbb{R})$, are the Borel and Lebesgue σ-algebras of $\mathbb{R}$, respectively, then*

$$\mathcal{B}(\mathbb{R})^{\otimes n} = \mathcal{B}(\mathbb{R}^n),$$

but

$$\mathcal{L}(\mathbb{R})^{\otimes n} \subsetneq \mathcal{L}(\mathbb{R}^n).$$

Proof. $\mathcal{B}(\mathbb{R}^n)$ is generated by definition by the open sets in $\mathbb{R}^n$, which are by definition at most countable unions of products of open sets. Thus, $\mathcal{B}(\mathbb{R}^n) \subset \mathcal{B}(\mathbb{R})^{\otimes n}$. Conversely, a rectangle $A \times B$ can be written as the intersection $A \times X \cap X \times B$, where both sets are measurable since the projection maps are continuous. We thus conclude that $\mathcal{B}(\mathbb{R})^{\otimes n} \subset \mathcal{B}(\mathbb{R}^n)$ (and an analogous argument holds for the Lebesgue measure). Finally, if $V \subset \mathbb{R}$ is unmeasurable, $V \times \{0\}$ is Lebesgue in $\mathbb{R}^n$ since it is null, but since the inclusion $i : x \to (x, 0)$ is measurable, it is not in the product σ-algebra. □

Remark 3.8.2. In general, since the projection maps are continuous in the product topology, they are measurable. Thus, any cylinder set is a finite intersection of preimages under projections of Borel sets, and is thus Borel measurable on the product space, which shows that $\bigotimes \mathcal{B}(X_i) \subset \mathcal{B}(\prod X_i)$. Conversely, note that **not every open set is a cylinder set** (since one can take unions of cylinder sets), so equality does not necessarily hold. However, if each space is second-countable (i.e. has a countable base) and the product is countable, then each open set is a union of cylinder sets, so $\mathcal{B}(\prod X_i) \subset \bigotimes \mathcal{B}(X_i)$.

Example 3.8.1. Take $X \times X$ with the discrete topology, where X has cardinality greater than $\mathbb{R}$. Then, the diagonal is a union of open sets and is therefore measurable. If the diagonal was measurable in the product σ-algebra, it has to be the union of at most uncountably many sets, so at least one set has two points $(u,u),(v,v)$, implying (u,v) is in the diagonal, a contradiction.

A popular type of question is to show that a certain subset of a set is Borel, i.e. a countable union and intersection of open or closed sets. The typical approach in these problems is to convert between logic-based definitions and the corresponding countable unions and intersections of open/closed sets.

Lemma 3.8.1. *The set of points of continuity of a measurable function $f : \mathbb{R} \to \mathbb{R}$ is G_δ.*

Proof. First, since we want to avoid talking about individual continuity points (as there are uncountably many of them), we rephrase continuity as a property on an open set. In particular, notice that f is continuous at c iff

$$\forall \epsilon > 0 \exists \delta > 0, \forall y, z \in (c-\delta, c+\delta) \implies |f(y) - f(z)| < \epsilon,$$

We thus define

$$A_n = \left\{ x : \exists \delta > 0, \forall y, z \in (x-\delta, x+\delta) \implies |f(y) - f(z)| < \frac{1}{n} \right\}.$$

Then,

$$\bigcap_n A_n = \{x : \forall \epsilon > 0 \exists \delta > 0, \forall y, z \in (x-\delta, x+\delta) \implies |f(y) - f(z)| < \epsilon\},$$

which is in fact the set $C(f)$ of continuity points of f. It remains to show that A_n is open. But if y is sufficiently close to $x \in A_n$, taking $\delta_y = \frac{\delta_x}{100}$ suffices, and we are done. □

Remark 3.8.3. This proof immediately generalizes to $f : X \to \mathbb{R}$ for an arbitrary metric space $\mathbb{R}$.

Lemma 3.8.2. *On $\mathbb{R}$, the converse holds - for every G_δ set $X \subset \mathbb{R}$, there exists a function that is continuous precisely on X.*

Proof. Let $X = \bigcap_i U_i$ be a G_δ set. WLOG, suppose that $U_{i+1} \subsetneq U_i$. First, note that the function $\chi_{U_i} + \frac{1}{2}\chi_{U_i^c \cap \mathbb{Q}}$ is continuous on the open set U_i, and discontinuous everywhere else, since if it was continuous at $x \in U_i^c$, it would be close to 0 or $\frac{1}{2}$ in some open neighborhood of x, and so it would have to be 0 or $\frac{1}{2}$ on some open neighborhood of x. But that is impossible since f is 0 or $\frac{1}{2}$ on a subset of the rationals/irrationals. We now define the function

$$f(x) := \frac{1}{2} \sum_{i=0}^{\infty} \frac{1}{2^i} (\chi_{U_i} + \frac{1}{2}\chi_{U_i^c \cap \mathbb{Q}}).$$

Note that the function takes the following values:

$$f(x) = \begin{cases} 1 & x \in \bigcap_i U_i, \\ 1 - 2^{-n-1} - 2^{-n-2} & x \in (U_n \backslash U_{n+1}) \cap \mathbb{Q}, \\ 1 - 2^{-n-1} & x \in (U_n \backslash U_{n+1}) \cap \mathbb{Q}^c, \\ \frac{1}{2} & x \in U_0^c \cap \mathbb{Q}, \\ 0 & x \in U_0^c \cap \mathbb{Q}^c. \end{cases}$$

Now, if $x \in \bigcap_i U_i$, for any $x_k \to x$, $x_k \in U_n$ for any large enough n, so

$$f(x_k) \geqslant 1 - 2^{-n-1} - 2^{-n-2} \to 1$$

as $x_k \to x$ and $n \to \infty$, so f is continuous on X. If $x \in U_0^c \cap \mathbb{Q}^c$, then f is continuous at x iff f is 0 on an open neighborhood of x, which is impossible since it is only 0 on a subset of irrationals. If $x \in U_0^c \cap \mathbb{Q}$, then f is continuous at x iff all irrational points around it are in $U_0 \backslash U_1$ and all rational points are in U_0^c, which contradicts the fact that U_0 is open. Finally, if $x \in U_n \backslash U_{n+1}$ is rational, then it can be approximated by a sequence x_k of irrational points in at least U_n, so

$$f(x_k) \geqslant 1 - 2^{-n-1} > 1 - 2^{-n-1} - 2^{-n-2},$$

which shows that f is not continuous at x. If $x \in U_n \backslash U_{n+1}$ is irrational, then if x is a limit of point of U_{n+1}, there is a sequence $x_k \to x$ such that

$$f(x_k) \geqslant 1 - 2^{-n-2} - 2^{-n-3} \geqslant 1 - 2^{-n-1},$$

and if x is not a limit point of U_{n+1}, there exists a sequence x_k of rationals in $U_n \backslash U_{n+1}$ converging to x, so that

$$f(x_k) = 1 - 2^{-n-1} - 2^{-n-2} < f(x),$$

which shows f is not continuous at x. Thus, the constructed example is continuous precisely on the set X.

□

Corollary 3.8.1. *There does not exists a function continuous precisely on* $\mathbb{Q}$.

We now consider a much more involved question of differentiability with this technique.

Proposition 3.8.2. *The set of points where a measurable* $f : \mathbb{R} \to \mathbb{R}$ *does not possess a finite derivative is* $G_{\delta\sigma}$*, i.e. the set of points of differentiability is* $F_{\sigma\delta}$.

Proof. Similarly, we rephrase nondifferentiability as a local property. Define $G(a, b) = \frac{f(a)-f(b)}{a-b}$. Then, notice that a function is not differentiable at c iff $\lim_{a \to c} G(a, c)$ does not exist, i.e.

$$\exists \epsilon > 0, \forall \delta > 0, \exists y, z \in (c - \delta, c + \delta) \implies |G(y, c) - G(z, c)| < \epsilon.$$

We thus define

$$A_{n,m} = \left\{ x : \exists y, z \in \left(x - \frac{1}{n}, x + \frac{1}{n} \right) \implies |G(y, x) - G(z, x))| > \frac{1}{m} \right\}.$$

We claim this is an open set. Indeed, if $|x' - x| < \delta := \frac{1}{n} - \max(|z - x|, |y - x|)$, where y, z are chosen to satisfy the condition above, then for $|x' - x| < \frac{\delta}{10}$,

$$|G(y, x') - G(z, x')| > \frac{1}{m}$$

and $y, z \in \left(x' - \frac{1}{n}, x' + \frac{1}{n}\right)$. Thus,

$$\bigcup_n \bigcap_m A_{n,m} = \{x : \forall \epsilon > 0 \exists \delta > 0 \exists y, z \in (x - \delta, x + \delta) : |G(y, x) - G(z, x)| > \epsilon\}$$

is precisely the set of points where f is not differentiable, and since $A_{n,m}$ is open, this is a $G_{\delta\sigma}$ set. Thus, the set of points $\Delta(f)$ of differentiability is $G^c_{\delta\sigma} = F_{\sigma\delta}$. □

Corollary 3.8.2. *There is no function that is differentiable precisely on the Vitali set.*

Corollary 3.8.3. *The set of points where the derivative is continuous is a G_δ subset of an $F_{\sigma\delta}$ set, i.e. a countable intersection of sets that are themselves an interection of an open set with an $F_{\sigma\delta}$ set, i.e. $C(f')$ is also an $F_{\sigma\delta}$ set. Recursively, this implies that $C(f^{(n)}), \Delta^n(f)$ are at worst $F_{(\sigma\delta)^n}$ sets.*

In fact, one can say something stronger about $C(f')$ for an everywhere differentiable function f.

Proposition 3.8.3. *If f is differentiable, then $C(f')$ is a dense G_δ subset of $\mathbb{R}$.*

Proof. The previous corollary implies that the set is G_δ. To show that it is dense, note that f' is the pointwise limit of continuous functions $f_n(x) = \frac{f(x+\frac{1}{n})-f(x)}{\frac{1}{n}}$. Define

$$D_n(f) = \{x : \limsup_{y \to x} f(y) - \liminf_{y \to x} f(y) \geqslant \frac{1}{n}\},$$

and note that since $C(f') = \bigcap D^c_n(f')$, it suffices to show that D^c_n is dense for all n, as $C(f')$ is then an intersection of dense open sets and therefore dense by the Baire category theorem.

For sake of contradiction, suppose $D^c_n \cap I = \varnothing$ for an open set I. Now, define $E_k = \bigcap_{i,j \geqslant k} \{x : |f_i(x) - f_j(x)| \leqslant \frac{1}{4n}\}$, and note that since f_n converge pointwise, i.e. $\bigcup_k E_k = X$, and E_k are closed, by Baire Category there exists k s.t. $E_k \cap \overline{I}$ has nonempty interior. But since each f_i is continuous, taking the limit $i \to \infty$ implies that on that interior, $|f(x) - f(y)| < \frac{3}{4n} < \frac{1}{n}$, contradicting the fact that $D^c_n \cap I = \varnothing$. □

Corollary 3.8.4. *If f is everywhere differentiable, then the set of points where f' is continuous is uncountable.*

Remark 3.8.4. Despite this, there exist everywhere differentiable functions whose derivatives are continuous only on a measure zero set. For example, consider the construction of the **Volterra function:** take the function $x^2 \sin(\frac{1}{x})$ on $[0, \alpha]$, cut it off at the largest value of x where the derivative is zero, mirror across $x = \alpha$, and extend it to be a constant. Then, translate it to the first interval removed from the fat Cantor set C_β of measure $0 < \beta < 1$. Repeat this process for each of the subintervals removed in the fat Cantor set. Then, the resulting function V_α has the following properties: since the complement of C_β is open and dense, the function is in fact differentiable with bounded derivative on C^c_β. On the fat Cantor set, the function identically vanishes, and is in fact differentiable with derivative 0, since $x^2 \sin(\frac{1}{x})$ is differentiable with derivative 0 at $x = 0$. However, V'_α is discontinuous on C_β, since there is a sequence of endpoints (at which the derivative is discontinuous) converging to every point of C_β. In particular, V'_α is not Riemann integrable. Now, in each of the interval removed, we can put another copy of the Volterra function, e.g. if V'_α is discontinuous on a fat Cantor set of measure $\frac{1}{2}$, we can cover half of the remaining measure by copies of the function. Repeating this countable process countably many times yields the Volterra function $V(x)$, which is differentiable everywhere but the derivative is discontinuous a.e. In particular, this a counterexample to the Fundamental Theorem of Calculus if the derivative is not assumed to be integrable.

3.8.1 Exercises

Problem 3.8.1 (Problem 1 Fall 2014). Show that $L^3(\mathbb{R}) \cap L^2(\mathbb{R})$ is Borel in $L^3(\mathbb{R})$.

Proof. The idea is to break up the problem into smaller pieces that are more manageable and where we have known results. In particular, note that we have a relation between the L^2 and L^3 norm on a finite measure set A, stating that $\|f\|_2 \lesssim \|f\|_3$. I claim that

$$A_{n,k} := \{f \in L^3 : \|f\|^2_{L^2([-k,k])} < n\}.$$

is open in L^3. Indeed, suppose $f \in L^3$ and $\|f\|_2^2 < n$. Hölder implies that for all $\epsilon > 0$, there exists a $\delta > 0$ such that $\|f - g\|_2 < \epsilon$ whenever $\|f - g\|_3 < \delta$. In particular, this implies that for any values of $\|f\|_2^2 < n$, we can choose $\delta > 0$ such that $\|g\|_2^2 \leqslant (\|f - g\|_2 + \|f\|_2)^2 < n$ whenever $\|f - g\|_3 < \delta$, showing that $A_{n,k}$ is open in L^3. Then,

$$L^3(\mathbb{R}) \cap L^2(\mathbb{R}) = \bigcup_{n\in\mathbb{N}} \bigcap_{k\in\mathbb{N}} A_{n,k},$$

i.e. $L^3 \cap L^2$ is a $G_{\delta\sigma}$ set. □

Problem 3.8.2. Show that $L^2(\mathbb{R}) \cap L^\infty(\mathbb{R})$ is Borel in $L^\infty(\mathbb{R})$.

Proof. The issue with applying the same method as for the last problem is that the Hölder bound gives us the reverse direction, i.e. $\|f\|_2 \lesssim \|f\|_\infty$. Thus, we instead attempt to prove that the sets

$$B_{n,k} = \{f \in L^\infty : \|f\|^2_{L^2([-k,k])} \leqslant n\}$$

are closed in L^∞. Indeed, suppose $f_n \in B_{N,k}$, $f_n \to f$ in L^∞. Then, by Fatou, $\|f\|_2^2 \leqslant \liminf \|f_n\|_2 \leqslant N$, and so $f \in B_{N,k}$, i.e. $B_{N,k}$ is closed in L^∞. Then, as before,

$$L^2(\mathbb{R}) \cap L^\infty(\mathbb{R}) = \bigcup_{n\in\mathbb{N}} \bigcap_{k\in\mathbb{N}} B_{n,k},$$

i.e. it is a F_σ set. □

Problem 3.8.3. Let $f : \mathbb{R}^n \to \mathbb{R}$ be measurable. Show that the set of points of continuity of f is Borel.

Proof. Again, the idea is to construct an intersection or union of open/closed sets. Here, one has to specifically rely on the fact that contuinuity is "local": it tells you something about the oscillation of a function in a neighborhood of a point. Particularly, we will use the following definition of continuity, which avoids talking about continuity "at a point."

$$f \text{ continuous at } x \iff \forall\epsilon > 0 \exists\delta > 0, |y - x| < \delta, |z - x| < \delta \implies |f(y) - f(z)| < \epsilon.$$

Then, define

$$C_\epsilon := \{x : \exists\delta > 0 : |y - x| < \delta, |z - x| < \delta \implies |f(y) - f(z)| < \epsilon\}.$$

Note that C_ϵ does not parametrize by δ, as that would be too weak to state openness or closedness. I claim that c_ϵ is open. Indeed, if $|x' - x| < \frac{\delta}{2}$, then $|y - x'|, |z - x'| < \frac{\delta}{2}$ satisfies the conditions of C_ϵ. Thus, C_ϵ is open, and

$$C(f) = \bigcap_{n \in \mathbb{N}} C_{\frac{1}{n}},$$

i.e. $C(f)$ is a G_δ set. □

Problem 3.8.4. Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuous. Show that the set of points of differentiability of f is Borel.

Proof. We attempt the same approach, yet as always, we want to use the right definition of differentiability. In particular, let us rely on the previous result, defining the continuous function $F(x, h) = \frac{f(x+h) - f(x)}{h}$.

$$f \text{ differentiable at } x \iff \forall \epsilon \exists \delta \exists Y : |h| < \delta \implies |F(x, h) - Y| < \epsilon.$$

Now, define

$$D_{\epsilon,\delta,Y} = \{x : |h| < \delta \implies |F(x, h) - Y| < \epsilon.\}$$

Then,

$$D(f) = \bigcup_{k \in \mathbb{Q}} \bigcap_{m \in \mathbb{N}} \bigcup_{n \in \mathbb{N}} D_{\frac{1}{m}, \frac{1}{m}, k},$$

which is a $G_{\delta\sigma\delta}$ set. For general measurable functions, note that $D(f) \subset C(f)$, so $D(f)$ is Borel is Borel on $C(f)$, i.e. it is Borel in $\mathbb{R}$. □

Problem 3.8.5. Let $T : C_c(\mathbb{R}) \to C_c(\mathbb{R})$ be such that $\|Tf\|_\infty \leqslant \|f\|_\infty$ and $\mu\{x : |Tf(x)| > \lambda\} < \frac{\|f\|_1}{\lambda}\}$. Show that $\|Tf\|_p \lesssim \|f\|_p$ for all $1 \leqslant p \leqslant \infty$.

3.9 Sobolev Spaces

We have seen that often times, we are able to take derivatives of functions in some sense. For instance, BV functions are precisely those whose derivative is a measure. But what if the derivative itself is a function?

Definition 3.9.1. Let $\Omega \subset \mathbb{R}^n$ be open. A function $g : \Omega \to \mathbb{R}$ is the α-th **weak derivative** of f if $\int_\Omega f D^\alpha \phi = (-1)^{|\alpha|} \int_\Omega D^\alpha g \phi$ for all $\phi \in C_c^\infty(\Omega)$.

We now consider the spaces of functions whose weak derivative is integrable.

Definition 3.9.2. For $k \in \mathbb{N}$ and $1 \leqslant p \leqslant \infty$ The **Sobolev space** $W^{k,p}(\Omega)$ is the space of functions whose first k derivatives are in L^p, i.e. the space of functions with finite norm

$$\|f\|_{W^{k,p}} = \left(\sum_{|\alpha| \leqslant k} \|D^\alpha f\|_p^p \right)^{\frac{1}{p}}$$

for $p < \infty$ and

$$\|f\|_{W^{k,\infty}} = \sum_{|\alpha| \leqslant k} \|D^\alpha\|_\infty.$$

Proposition 3.9.1. *$W^{k,p}(\Omega)$ is a Banach space.*

Proof. If f_n is Cauchy is the $W^{k,p}$ norm, then the candidates for the limit f and its derivatives $D^\alpha f$ are clear. It remains to show that

$$\int f D^\alpha \phi = \lim_{n\to\infty} \int f_n D^\alpha \phi = (-1)^{|\alpha|} \lim_{n\to\infty} \int D^\alpha f_n \phi = (-1)^{|\alpha|} \int D^\alpha f \phi,$$

which shows the claim. The case $p = \infty$ can be handled similarly. □

Since $W^{k,p}$ can be thought of as a collection of L^p spaces, Sobolev spaces inherit many properties from L^p. For instance, $W^{k,p}$ is separable iff $p < \infty$, reflexive with dual $W^{k,p'}$ for $1 \leqslant p < \infty$, and $W^{k,2} := H^k$ is a Hilbert spaces. As always, we have density results, such as the following:

Proposition 3.9.2. *For $p < \infty$, C_c^∞ is dense in $W^{k,p}$ in the $W^{k,p}$ norm.*

Proof. Take an approximation ϕ_ϵ to the identity. Then, since $D^\alpha(f * \phi_\epsilon) = (D^\alpha f) * \phi_\epsilon$, if $\Phi \in C_c^\infty$ is 1 on $B(0,1)$ and 0 on $B(0,2)$, then $\Phi(\epsilon x)(f * \phi_\epsilon)$ is a sequence of compactly supported functions converging to f in $W^{k,p}$. □

But really, why do we care about Sobolev spaces? This is true mainly because of a wide number of so-called Sobolev embedding theorems, which trade regularity for higher integrability. We begin with the simplest case of $W^{1,p}$.

Definition 3.9.3. For $p < n$, define the **Sobolev conjugate** $p^* > p$ of $W^{k,p}$ by

$$\frac{1}{p^*} = \frac{1}{p} - \frac{k}{n}.$$

Theorem 3.9.1 (Gagliardo-Nirenberg)**.** *If Ω is bounded, then $W^{1,p}(\Omega) \hookrightarrow L^{p^*}(\Omega)$ is a continuous embedding, i.e.*

$$\|u\|_{L^{p^*}} \leqslant C\|u\|_{W^{1,p}}.$$

Proof. Note that

$$|u| = \left|\int u_{y_i} dy_i\right| \leqslant \int |\nabla u| dy_i.$$

Thus, by generalized Hölder,

$$\int |u|^{\frac{n}{n-1}} dx_1 \leqslant \int \prod_i \left(\int |\nabla u| dy_i\right)^{\frac{1}{n-1}} dx_1 \leqslant \left(\int |\nabla u| dy_1\right)^{\frac{1}{n-1}} \left(\prod_{i=2}^n \iint |\nabla u| dx_1 dy_i\right)^{\frac{1}{n-1}}.$$

Next, we pull out the y_2 factor, integrate with respect to x_2, and use generalized Hölder again to get

$$\int |u|^{\frac{n}{n-1}} dx_1 dx_2 \leqslant \left(\iint |\nabla u| dx_1 dy_2\right)^{\frac{2}{n-1}} \prod_{i\geqslant 3} \left(\iiint |\nabla u| dx_1 dx_2 dy_i\right)^{\frac{1}{n-1}}.$$

Continuing in this fashion, we get

$$\int |u|^{\frac{n}{n-1}} dx \leqslant \left(\int |Du| dx\right)^{\frac{n}{n-1}},$$

which yields the estimate for $p = 1$. For $p > 1$, apply the estimate to $v = |u|^\gamma$ and use Hölder to get that if γ is chosen so that

$$\frac{\gamma n}{n-1} = (\gamma - 1)\frac{p}{p-1},$$

then

$$\frac{\gamma n}{n-1} = \frac{np}{n-p} = p^*.$$

This yields the general Gagliardo-Nirenberg inequality. □

The way to interpret Gagliardo-Nirenberg is that for one derivative in L^p, you get an extra $\frac{1}{n}$ in integrability.

Now, what happens if $p > n$? Surprisingly, it then turns out that the function is Hölder continuous.

Theorem 3.9.2 (Morrey's Inequality)**.** *The inclusion $W^{1,p}(\mathbb{R}^n) \to C^{0,\gamma}(\mathbb{R}^n)$ is continuous for $\gamma = 1 - \frac{n}{p}$.*

Proof. □

Thus, we have that for a function with one derivative in L^p, we get $\frac{1}{n}$ of integrability if $p < n$ and $1 - \frac{n}{p}$ of Hölder continuity if $p > n$. We can now generalize these statements to $W^{k,p}$.

We may now generalize this to having k derivatives in L^p, which is known as the general Sobolev embedding theorem.

Theorem 3.9.3 (General Sobolev Inequalities)**.** *(a)* ***Sobolev Embedding Theorem:*** *Let p^*, q^* be the Sobolev conjugates of $W^{k,p}, W^{l,q}$, respectively. Then, if $p < n$ and $q > p, k > l$, if*

$$p^* = q^*,$$

then one has a continuous embedding

$$W^{k,p}(\mathbb{R}^n) \subseteq W^{l,q}(\mathbb{R}^n).$$

(b) ***Rellich-Kondrachov Theorem:*** *If $k > l$ and*

$$p^* > q^*,$$

then on a bounded open set U, then the embedding is compact.

(c) If $pk > n, r \in \mathbb{N}$, and

$$r = \lfloor \frac{pk-n}{p} \rfloor, \alpha = \left\{ \frac{pk-n}{p} \right\},$$

then one has a continuous embedding

$$W^{k,p}(\mathbb{R}^n) \subset C^{r,\alpha}(\mathbb{R}^n).$$

Remark 3.9.1. How do we interpret the inequality above? First of all, if we embed a Sobolev space into another Sobolev space, we will lose derivatives and gain integrability, so $q > p, k > l$. Then, we need $p < n$ as in the proof of Gagliardo-Nirenberg. The case $p^* = q^*$ represents a critical case of the inequality, and the if the two are not equal, then we have enough regularity to establish compactness. Finally, the last part is a generalization of Morrey's inequality, which tells us that if our derivatives are very integrable, then that is as good as the function being continuous differentiable.

Corollary 3.9.1. *If $pk > n$, then $W^{k,p}(\mathbb{R}^n) \subset C(\mathbb{R}^n)$, and thus consists of continuous functions. For example, for $n = 1$, $W^{1,1}(\mathbb{R}) = AC(\mathbb{R})$ is the space of continuous functions. For $n = 2$, $H^2(\mathbb{R}^2) \subset C(\mathbb{R}^2)$.*

Here is the proof of the Rellich-Kondrachov Theorem:

Proof. Let $u_m \in W^{k,p}$ be a bounded sequence, and let u_m^ϵ be the corresponding mollifiers. The goal is to show that $u_m^\epsilon \to u_m$ uniformly in m as $\epsilon \to 0$, as then, by applying Arzela-Ascoli on u_m^ϵ, one may obtain

$$\limsup \|u_m^\epsilon - u_n^\epsilon\|_q = 0, \|u_n^\epsilon - u_n\|_q \leqslant \frac{\delta}{2} \implies \limsup \|u_m - u_n\| \leqslant \delta,$$

and finish with standard diagonal argument. To show uniform convergence, we can easily show that $u_m^\epsilon \to u_m$ uniformly in L^1, and since $q^* < p^*$, by interpolation, we can bound the L^{q^*} norm by L^1 and L^{p^*} norms, where the latter may then be bounded by Gagliardo-Nirenberg. □

Now, recall that the Fourier symbol of the derivative operator is ξ^1. What if we take noninteger powers of ξ?

Definition 3.9.4. For $k \in \mathbb{R}, p < \infty$, the **Bessel potential space** $H^{k,p}(\Omega)$ is the space of functions with finite Sobolev norm

$$\|f\|_{H^k(\Omega)} = \|\mathcal{F}^{-1}\langle\xi\rangle^k \mathcal{F} f\|_p,$$

where $\langle\xi\rangle = (1 + |\xi|^2)^{\frac{1}{2}}$ is the **Japanese bracket.**

Finally, we may attempt to generalize the Hölder condition to the L^p setting to attain yet another possible definition.

Definition 3.9.5. For $s \in (0,1)$, define the **Sobolev-Slobodeckij space** $W^{s,p}$ as the space of functions with finite norm

$$\|u\|_{W^{s,p}} = \left(\int |u|^p + \int \frac{|u(x) - u(y)|^p}{|x-y|^{n+sp}} dxdy\right)^{\frac{1}{p}}.$$

Define $W^{k,p}$ for $k > 1$ as the sap

Turns out that such spaces are equivalent to $W^{k,p}$.

Proposition 3.9.3. *$H^{k,p} = W^{k,p}$ whenever k is a nonnegative integer.*

Proof. It suffices to show equivalence of norms. Suppose $f \in W^{k,p}$. We appeal to the Mikhlin multiplier theorem, which says that if m is a smooth bounded function s.t. $|x|^k|\nabla^k m|$ is bounded for $0 \leqslant k \leqslant \frac{n}{2} + 1$, then m is an L^p multiplier. □

3.10 The Laplacian: A Case Study

One of the most important linear operators in analysis is the Laplacian operator $-\Delta$, which represents the negative sum of the second partial derivatives of a function. We do a brief, yet in depth summary of the operator and its spectral and analytic properties.

Since not every function is differentiable, we first want to clearly define the domain of $-\Delta$. Since we want an inner product structure, for now we consider $-\Delta : A \subset L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$. Natural choices of domain are $A = C^k(\mathbb{R}^n), H^2(\mathbb{R}^n)$, for $k \geqslant 2$. The next proposition shows that one of these is considerably more natural than the others.

Lemma 3.10.1. *For $A = C_0^\infty(\mathbb{R}^n)$, $-\Delta$ is closable with the closure $(-\Delta, H^2(\mathbb{R}^n))$.*

Proof. It suffices to show that for $(f_n, -\Delta f_n) \to (0, g)$ in L^2, one has $g = 0$. By Fourier transforms, $\|g\|_2 = \lim_{n\to\infty} \||\xi|^2 \widehat{f_n}\|_2$ and $\|\widehat{f_n}\|_2 \to 0$. But we in fact know that $\widehat{f}_n$ is in the Schwartz space, so we immediately see that $\|g\|_2 = 0$. To see that the closure contains H^2, we note that C_0^∞ is dense in H^2 in the H^2-norm. Finally, the fact that $H^2(\mathbb{R}^n)$ is closed as a domain follows from the fact that if $f_n \to f, -\Delta f_n \to g \in L^2$, then on the Fourier side, $|\xi|^2 \widehat{f_n} \to |\xi|^2 \widehat{f} = \widehat{g}$ in L^2 (as can be easily checked on compact subsets) and the claim follows. ☐

We now want to show that the Laplacian is self-adjoint. But in fact, our definition of self-adjointness is quite tricky to demonstrate, so we need a simpler criterion first.

Lemma 3.10.2 (Criterion for Self-Adjointness)**.** *A closed symmetric operator T is self-adjoint iff $\ker(T^* \pm i) = 0$ iff $\mathrm{Im}(T \pm i) = H$.*

Proof. Note that a closed operator has closed kernel and image, so the equivalence of the last two statements is equivalent by the orthogonal decomposition of a Hilbert space. To show $D(T^*) \subset D(T)$, let $f \in D(T^*), \phi = (T^* + i)f$, and $g \in D(T)$ s.t. $\phi = (T + i)g$. Then, since $T^* g = Tg$, we get $(T^* + i)(f - g) = 0$, i.e. $f = g \in D(T)$. ☐

Corollary 3.10.1. *A closed symmetric operator T is self-adjoint iff $\sigma(T) \subset \mathbb{R}$.*

Remark 3.10.1. The same proof applies on a bounded open subset of $\mathbb{R}^n$.

Proposition 3.10.1. *The Laplace operator is self-adjoint $-\Delta : H^2(\mathbb{R}^n) \subset L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ with essential spectrum $\sigma(-\Delta) = \sigma_{ess}(-\Delta) = [0, \infty)$.*

Proof. Notice that the Laplace operator is the Fourier multiplier of $|\xi|^2$. We borrow from the theory of multiplication operators, which states that the spectrum of a multiplication operator is its essential range, i.e. the support of the pushforward measure $f_*\mu$, where eigenvalues λ are s.t. $\mu(f = \lambda) > 0$. Since the range of $|\xi|^2$ is $[0, \infty)$, so is the spectrum of $-\Delta$, which is easily seen to be purely continuous. Moreover, since the spectrum is real and the operator is easily seen to be symmetric, the operator is self-adjoint. ☐

Remark 3.10.2. This method easily allows us to construct eigenfunctions any (generalized) function supported on $\{x : f = \lambda\}$ is the Fourier transform of an eigenfunction. For instance, we identify the plane waves $e^{i\lambda x}$ as "eigenfunctions" of the Laplacian with eigenvalues λ.

We can now generalize our approach to the Schrodinger operator $-\Delta + V(x) : H^2(\mathbb{R}^n) \subset L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$. In general V may not be bounded, so the multiplication operator of V might be an unbounded operator.

Proposition 3.10.2. *$-\Delta + V$ essentially self-adjoint on $H^2 \cap D(V)$ for $V \in L^2_{loc}, V \geqslant 0$.*

Proof. The same argument as before shows that the operator is closable on C_c^∞. The same argument as before shows that H^2 is contained in the closure. Finally, the same argument as before shows that $-\Delta + V$ is closed on L^2.

For multiplication by V to be symmetric, V clearly has to be real-valued. Finally, we get that the spectrum of the operator depends on $\widehat{V}$ - namely, $\sigma_{ess}(-\Delta + V)$ is the essential range of $-\Delta + V$, and eigenvalues are values λ where $\mu(|\xi|^2 + \widehat{V} = \lambda) > 0$. ☐

3.11 Ergodic Theory

Several questions on the qualifying exam pertain to ergodic theory. We briefly review the main facts here.

Definition 3.11.1. A **measure-preserving transformation** $T : X \to X$ on a probability space is a map such that $\mu(T^{-1}(A)) = \mu(A)$ for all measurable A.

Definition 3.11.2. A measure-preserving transformation is said to be **ergodic** if $\mu(T^{-1}(E)\Delta E) = 0$ implies $\mu(E) = 0$ or $\mu(E) = 1$.

Lemma 3.11.1. *For* $\alpha \in \mathbb{R}\backslash\mathbb{Q}$, $\{\alpha n \mod 1 : n \in \mathbb{N}\}$ *is dense in* $[0, 1]$.

Proof. By pigeonhole principle, subdividing $[0, 1]$ into N intervals, there are $j, k \in \mathbb{N}$ such that $(j - k)\alpha \mod 1 \in (-\frac{1}{N}, \frac{1}{N})$. The rest follows by adding this number to itself at most N times. ☐

In fact, there is a much stronger definition for a subset.

Definition 3.11.3. A sequence in $[0, 1]$ is said to **equidistributed** if $\lim_{n\to\infty} \mu(a_n \in [c, d]) = d - c$.

Proposition 3.11.1. *A sequence is equidistributed in* $[a, b]$ *iff* $\lim_{n\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(a_n) = \frac{1}{b-a}\int_a^b f(x)dx$ *in the Riemann integral sense.*

Proof. Equidistribution is equivalent to the Riemann integrability of indicator functions. Conversely, approximating f by step functions below and above by linearity yields Riemann integrability. ☐

Theorem 3.11.1 (Weyl's Equidistribution Theorem)**.** *A sequence is equidistributed in* $[0, 1]$ *iff*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} e^{2\pi i a_n} = 0.$$

Proof. If a sequence is equidistributed, thiis follows immediately by the Riemann integrability criterion. Conversely, if the criterion holds, it holds for every trigonometric polynomial, and by Stone-Weierstrass, for (almost) every continuous function. Then, approximating step functions by continuous functions as before, the proof concludes. ☐

Example 3.11.1. Let $\alpha \in \mathbb{R}\backslash\mathbb{Q}$. Then,

$$\frac{1}{N}\sum_{n=0}^{N} e^{2\pi i\alpha n} = \frac{1}{N(1 - e^{2\pi i\alpha})} \to 0,$$

so multiples of an irrational number are in fact equidistributed in $[0, 1]$.

The importance of ergodic theory are the so-called ergodic theorems, which state that for ergodic transformations, the average in space and in time are identical.

Theorem 3.11.2. *Let* $T : X \to X$ *be a measure-preserving transformation on a finite measure space and let* $f \in L^1$*. Define the time and space averages*

$$\hat{f}(x) = \lim_{n\to\infty} \frac{1}{N}\sum_{n=0}^{N} f(T^k x), \overline{f} = \frac{1}{\mu(X)}\int f d\mu.$$

Then $\hat{f} \in L^1$*, and if* T *is ergodic,* $\hat{f} = \overline{f}$ *a.e., with* $\int f d\mu = \int \hat{f} d\mu$ *and* $\hat{f} = \hat{f} \circ T$.

4 Complex Analysis

4.1 Holomorphic Functions

The following are the equivalent definitions of a holomorphic function:

Definition 4.1.1. (a) A holomorphic function $f : U \to \mathbb{C}$ is complex differentiable, i.e.

$$\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

exists.

(b) A holomorphic function $f : U \to \mathbb{C}$ is a complex function given locally by a power series $f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z - z_0)^n$ that that converges normally to f on $|z - z_0| < R$, where R is the smallest distance to where f is undefined.

(c) A holomorphic function is one satisfying the **Cauchy-Riemann (C-R)** equations: if $f = u+iv$, $u_x = v_y$ and $u_y = -v_x$.

To prove this, we first need to prove a fundamental result known as the Cauchy-Goursat theorem.

Theorem 4.1.1 (Goursat Theorem)**.** *If f is complex differentiable in the sense of (a) on an open region U, for any triangle Δ contained in U, $\int_\Delta f(z) = 0$.*

Proof. For contradiction, suppose not, i.e $|\int_\Delta f(z)dz| > \epsilon$. Iteratively subdivide the triangle into subtriangles, and by pigeonhole principle, pick a point z^* in a sequence of triangles where $|\int_{\Delta_n} f(z)dz| \geqslant \frac{\epsilon}{4^n}$. But since each triangle has half the diameter and perimeter of the previous one,

$$\left|\int f(z)dz - (f(z^*) + f'(z^*)(z - z^*))dz\right| \leqslant \epsilon' \int |z - z^*|dz \leqslant \epsilon' \text{diam}(\Delta_n)|\Delta_n| = \frac{\epsilon' \text{diam}(\Delta_0)|\Delta_0|}{4^n},$$

which is a contradiction for small enough ϵ'. □

Corollary 4.1.1. *Approximating an arbitrary simple curve by polygons, which are further subdivided into triangles, and approximating null homotopic curves by simple closed curves implies **Cauchy's Theorem:** for $f : U \to \mathbb{C}$ complex differentiable and γ a closed C^1 curve inside U homotopic to a point,*

$$\int_\gamma f(z)dz = 0.$$

Theorem 4.1.2. *(Cauchy's Integral Formula and Estimates) If f is holomorphic on U and γ is a circle of radius R in U,*

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \oint_\gamma \frac{f(z)}{(z - z_0)^{n+1}} dz$$

and

$$|f^{(n)}(z_0)| \leqslant \frac{n!M}{R^n}.$$

Proof. WLOG, suppose γ is a circle. For $n = 0$,

$$\left|\oint_\gamma \frac{f(z) - f(z_0) + f(z_0)}{z - z_0} dz - 2\pi i f(z_0)\right| \to 0,$$

as the left part is bounded (since f is differentiable) and the right part can be evaluate to be $2\pi i f(z_0)$. Then, by the geometric series formula

$$f(w) = \frac{1}{2\pi i}\oint_\gamma \sum_{n=0}^{\infty} \frac{f(z)}{(z-z_0)^{n+1}}(w-z_0)^n dz = \sum_{n=0}^{\infty}(w-z_0)^n \oint_\gamma \frac{f(z)}{(z-z_0)^{n+1}}dz,$$

where the change in integral and sum is justified by taking a small enough circle γ. □

Corollary 4.1.2. *This argument shows that a holomorphic function f has a power series expansion that is locally uniformly convergent in any circle that the singularities of f. Moreover, any formal power series is a sequence of holomorphic functions that converges locally uniformly, and thus defines a holomorphic function on its radius of convergence $R = \frac{1}{\limsup |a_n|^{\frac{1}{n}}}$. This, along with a simple calculation showing that the C-R equations are equivalent to complex differentiability, shows that the three definitions of a holomorphic function are equivalent.*

Remark 4.1.1. Introducing the **Wirtinger derivatives**

$$\partial_z = \frac{1}{2}(\partial_x - i\partial_y), \partial_{\overline{z}} = \frac{1}{2}(\partial_x + i\partial_y),$$

C-R implies that f is holomorphic iff $\partial_{\overline{z}} f = 0$, and $f'(z) = \partial_z f(z)$.

Remark 4.1.2. On the boundary of the disk of convergence, the power series for a holomorphic function f can converge in any way possible - absolutely, conditionally but not absolutely, or it may diverge at any subset of $\partial\mathbb{D}$. For example,

(a) $\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n$ diverges everywhere on the boundary.

(b) $-\log(1-z) = \sum_{n=0}^{\infty} \frac{z^n}{n}$ converges everywhere conditionally except at $z = 1$.

(c) $\sum_{n=1}^{\infty} \frac{z^n}{n^2}$ converges everywhere on the boundary absolutely.

If $f : U \to \mathbb{C}$ is holomorphic, we write $f \in H(U)$. One can check that a given function is holomorphic using Morera's theorem:

Theorem 4.1.3 (Morera's Theorem)**.** *If $f : U \to \mathbb{C}$ is continuous and $\int_\Delta f(z)dz = 0$ for every triangle $\Delta \subset U$, then $f \in H(U)$.*

Proof. Define $F(z_0) = \int_\gamma f(z)dz$, where γ starts at a fixed point a and ends at z_0. The conditions, along with a polygonal approximation argument, imply that F is complex differentiable and so is holomorphic. By the fundamental theorem of calculus, since $F' = f$ and analyticity of F, it follows that f is holomorphic. □

Corollary 4.1.3. *The proof shows that a holomorphic function f on **any** domain U has an antiderivative F iff $\oint_\gamma f = 0$. In particular, every holomorphic function locally has an antiderivative. The necessity of the zero integral condition follows from the fact that by the Fundamental Theorem of Calculus, $\oint f(z) = F(z_0) - F(z_0) = 0$.*

The following are main results and theorems that frequently appear on the analysis qual:

Theorem 4.1.4. *(Maximum, Minimum Modulus and Mean Value Formulae) If f is holomorphic on U and γ is a circle of radius R in U,*

$$f(z_0) = \frac{1}{2\pi i}\oint_\gamma \frac{f(z)}{z-z_0}dz = \frac{1}{2\pi}\int_0^{2\pi} f(z_0 + Re^{i\theta})d\theta = \frac{1}{\pi R^2}\oiint_{B(0,R)} f(z)dz = \frac{1}{\pi R^2}\int_0^{2\pi}\int_0^R f(re^{i\theta})rdrd\theta.$$

If $|f|$ attains a (local) maximum on U, f is constant. If $|f|$ is bounded below on U by a positive constant, then if $|f|$ attains a (local) minimum on U, f is constant. Moreover, if f is continuous on ∂U, $|f|$ attains its maximum and (if f has no zeros in U) minimum on ∂U.

Proof. Mean value formulae follow immediately from Cauchy's integral formula, and directly imply the maximum and minimum principles. □

When discussing convergence of complex functions, the most natural setting is that of locally uniform convergence:

Definition 4.1.2. A sequence $f_n : U \to \mathbb{C}$ converges **normally** to $f : U \to C$ if $f_n \to f$ uniformly on compact subsets of U.

Proposition 4.1.1. *If $f_n \in H(U)$ and $f_n \to f$ normally, $f \in H(U)$ and $f_n^{(k)} \to f^{(k)}$ normally.*

Proof. The first part follows from Morera's theorem. The second part follows from Cauchy's Integral Formula. □

The following is a fundamental characterization of holomorphic functions.

Theorem 4.1.5 (Open Mapping)**.** *Every nonconstant holomorphic map is open, i.e. the image of an open set is open.*

4.2 Exercises

Problem 4.2.1. If $f \in H(\mathbb{D})$ has a pole at $z = 1$, then the Taylor series for f diverges everywhere on the boundary.

Proof. If not, $\sum a_n z^n = c$ for some $|z| = 1$, so $a_n \to 0$, so considering the series $b_n = a_n - a_{n=1}$, we see that $\sum b_n z^n = (1-z)\sum a_n z^n = (1-z)f(z)$. But as $z \to 1$, by Abel's theorem the former converges to 0 and the latter cannot converge to 0 because of the pole, which is a contradiction. □

Problem 4.2.2 (Fall 2014 Problem 9). Let $\Omega \subset \mathbb{C}$ be an open connected set. If f_n is a sequence of injective holomorphic functions on Ω that converges normally to f, then, if f is nonconstant, f is injective.

Proof. Note that f is injective iff f' does not vanish on U. In particular, since f_n converges locally uniformly to f, f_n' also converges locally uniformly to f'. Suppose f' is not injective. Then, $f'(z_0) = 0$ for some $z_0 \in U$. In particular, by the argument principle, over a sufficiently small circle of radius γ around z_0 on which f' does not vanish, $\frac{1}{2\pi i}\int_\gamma \frac{f''}{f'} = 1$, but $\frac{1}{2\pi i}\int_\gamma \frac{f_n''}{f_n'} = 0$ for all n. But this is a contradiction, since $|f'|$ is bounded on γ, so for $\epsilon < \inf_\gamma |f'|$, for sufficiently large n one has that $\frac{f_n''}{f_n'} \to \frac{f''}{f'}$ uniformly, contradicting the difference in the integral values. □

Problem 4.2.3 (Spring 2014 Problem 9). Prove that if $f_n \to f$ normally on an open connected set $\Omega \subset \mathbb{C}$, f_n, f holomorphic, $f_n(z) \neq 0$, then either f is either identically zero or vanishes nowhere.

Proof. Essentially a special case of the last problem. □

4.3 Zeros and Poles

Lemma 4.3.1 (Isolated Zeros)**.** *The zeros of a nonzero holomorphic function have finite order and are isolated.*

Proof. Since a holomorphic function is defined locally by a power series, $f(z) = z^k g(z)$ where g is analytic and $g(0) \neq 0$. This shows that the zero has finite order and is isolated. □

Corollary 4.3.1 (Identity Lemma)**.** *If $f, g \in H(U)$ agree on a set with a limit point in U, $f = g$.*

Definition 4.3.1 (Poles)**.** A complex function with isolated singularities has three types of singularities:

(a) A singularity z_0 is **removable** if $\lim_{z\to z_0} f(z) = f(z_0)$.

(b) A singuarity z_0 is a **pole of order** k if $\lim_{z\to z_0}(z - z_0)^n |f(z)| = \infty$ for $0 \leqslant n \leqslant k - 1$ but not for $n = k$.

(c) An **essential singularity** if neither is true.

Lemma 4.3.2 (Riemann's Theorem on Removable Singularities)**.** *If z_0 is an isolated singularity and f is bounded in a neighborhood of z_0, z_0 is a removable singularity.*

Proof. Note that $f(z)(z-z_0)^2$ is holomorphic at z_0, since it has zero derivative. Thus, $f(z)(z-z_0)^2 = h(z)$, where $h(z_0) = h'(z_0) = 0$. We thus have that $f(z) = \frac{h(z)}{(z-z_0)^2} = a_2 + a_3(z - z_0) + ...$ □

Definition 4.3.2. A **meromorphic function** $f : U \to \mathbb{C}$ is a function holomorphic outside a discrete (that is, closed countable) set of poles.

Lemma 4.3.3. *Every meromorphic function on $\mathbb{C}$ is a ratio of two entire functions.*

Proof. Given a meromorphic function f with poles a_n, by Weierstrass's Theorem, there exists a function g that has zeros a_n. Then, $fg = h$ is holomorphic, so $h = \frac{f}{g}$. □

Here are three important results that explain the behavior of meromorphic functions near singularities:

(a) **Little Picard's Theorem:** The image of an entire function misses at most one point of $\mathbb{C}$.

(b) **Casorati-Weierstrass:** The image of a function in a neighborhood of an essential singularity is dense in $\mathbb{C}$.

(c) **Great Picard's Theorem:** In a neighborhood of an essential singularity, a holomorphic function takes on all except at most one value of $\mathbb{C}$ infinitely many times.

(d) **Generalized Great Picard:** Any holomorphic map $M\backslash\{w\} \to \mathbb{CP}$ attains all except **at most two** values of $\mathbb{CP}$ infinitely many times in any neighborhood of w.

Some important corollaries follow when considering the singularity of a function at ∞.

Proposition 4.3.1. *If f is entire and has a removable singularity at ∞, then f is constant. If f has a pole at ∞, then it is a polynomial.*

Proof. The first claim follows immediately from Liouville's theorem. For the second claim, we see that f extends to a meromorphic function on $\mathbb{CP}$, so it is rational and has no poles in $\mathbb{C}$. Thus, it is a polynomial. $\square$

Remark 4.3.1. If f is entire and a polynomial, then by the Fundamental Theorem of Algebra it attains every value. If not, then it has an essential singularity at ∞, so since f never attains the value ∞ of $\mathbb{CP}\backslash\{\infty\}$, we see that the generalized Great Picard theorem implies the Little Picard Theorem.

4.4 Infinite Products

Recall that by the Weierstrass M-Test, the power series for a holomorphic function converges normally on the disk of convergence. In latter sections, we are interested in considering the convergence of infinite sums and products for meromorphic functions.

Definition 4.4.1. Given a countable subset $X = \{a_n\} \subset \mathbb{C}$ and a branch of logarithm with a branch cut that avoids X, we say $\prod_{n=1}^{\infty} a_n$ converges iff $\sum_{n=1}^{\infty} \log(a_n)$ converges. If the sum converges to $-\infty$, the infinite product is said to diverge to 0.

Remark 4.4.1. Note that if $a_n \geqslant 0$ for all n, since

$$\sum a_n \leqslant \prod(1 + a_n) \leqslant e^{\sum a_n},$$

then $\prod_{n=1}^{\infty}(1 + a_n)$ converges iff $\sum_{n=1}^{\infty} a_n$ converges.

Corollary 4.4.1. *A product $\prod_{n=1}^{\infty}(1 + a_n)$ is said to **converge absolutely** if $\prod_{n=1}^{\infty}(1 + |a_n|)$ converges. Then, the remark implies that $\sum_n a_n$ converges absolutely if and only if $\prod_{n=1}^{\infty}(1 + a_n)$ converges absolutely.*

Lemma 4.4.1. *The Cauchy criterion for convergence of products is as follows: the product converges if for any ϵ there exists a K s.t. $|\prod_n^m a_k - 1| < \epsilon$ for $n, m \geqslant K$.*

Proposition 4.4.1. *If a product converges absolutely, then it converges. In particular, if $a_n \geqslant 0$, $\prod(1 - a_n)$ converges iff $\sum a_n$ converges.*

Proof. The first statement follows from the Cauchy criterion and the inequality

$$\left|\prod_n^m(1 + a_k) - 1\right| \leqslant \prod_n^m(1 + |a_k|) - 1.$$

Then, if $\sum a_n$ converges, $\prod(1 - a_n)$ converges absolutely and therefore converges. Conversely, if $\sum a_n$ diverges and $\prod(1 - a_n)$ converges, $a_n \to 0$ so since $(1 + a_n)(1 - a_n) \leqslant 1 - a_n^2 \geqslant 0$, we get that $\prod(1 - a_n)(1 + a_n) \leqslant 1$, so $\prod(1 - a_n) \to 0$, which is a contradiction. $\square$

Remark 4.4.2. The examples $a_n = \frac{1}{\sqrt{[\frac{n}{2}]}}$ (where the product diverges to 0, but the sum converges) and $a_{2n} = \frac{\sqrt{n+1}}{\sqrt{n}}, a_{2n+1} = \frac{\sqrt{n}}{\sqrt{n+1}}$ (where the product converges, but the sum diverges) show that when we have both negative and positive terms, the convergence of $\prod(1 + a_n)$ and $\sum a_n$ is unrelated.

Corollary 4.4.2. *An infinite product $\prod_{n=1}^{\infty}(1 + f_n)$ of holomorphic functions converges to a holomorphic function if $\sum_{n=1}^{\infty} |f_n|$ converges.*

4.5 Weierstrass, Hadamard, Laurent, Mittag-Leffler, Jensen

We are often times interested in seris/product expansions for holomorphic/meromorphic functions. Their existence is provided through the following theorems:

Theorem 4.5.1 (Laurent Series)**.** *There exists a unique annulus at z_0 on which a function f with an isolated singularity at z_0 has a **Laurent series** of the form*

$$f(z) = \sum_{n=-\infty}^{\infty} a_n(z - z_0)^n,$$

*where $a_n = \frac{1}{2\pi i} \oint \frac{f(z)}{(z-z_0)^{n+1}} dz$ that converges normally on the annulus. If infinitely many negative terms of the Laurent series are nonzero, f has an essential singularity at z_0. Otherwise, if a_{-m} is the first nonzero coefficient, f has a pole of order m, and if the expansion has no negative terms, f has a removable singularity. Moreover, $R^{\pm} = \frac{1}{\limsup |a_{\pm n}|^{\frac{1}{n}}}$ are the inner and outer radii of the annulus. $\sum_{-\infty}^{-1} a_n z^n$ is called the **principal part** of f. Moreover, if if f is a holomorphic function in an annulus, then its Laurent series converges normally to f on that annulus.*

Proof. First, consider a formal Laurent series with $R^{\pm}$ defined as in the proof. Then, by the Weierstrass M-Test, the partial sums converge locally uniformly and thus define a holomorphic function on the annulus $R^- < |z - z_0| < R^+$. Conversely, if f is a holomorphic function in the annulus, one applies Cauchy's formula on the inner part of the annulus and the upper part of the annulus, on the intersection, f is given by its Laurent series, with the partial sums converging normally. □

Example 4.5.1. Consider the function $e^{\frac{1}{z}} - \frac{1}{z-2}$. This function has two Laurent series - one in the annulus $0 < |z| < 2$, given by

$$\sum_{n=0}^{\infty} \frac{z^{-n}}{n!} + \frac{1}{2} \sum_{n=0}^{\infty} \left(\frac{z}{2}\right)^n,$$

and one in the annulus $|z| > 2$, given by the Laurent series for $e^z - \frac{z}{1-2z}$ at 0, which is

$$\sum_{n=0}^{\infty} \frac{z^{-n}}{n!} - z \sum_{n=0}^{\infty} (2z)^n.$$

Arguably the most fundamental result regarding meromorphic functions is that of the residue theorem:

Definition 4.5.1. Let f be meromorphic, and z_0 be a singularity of f. $\text{Res}(f, z_0) = a_{-1} = \int_\gamma f(z)dz$ is called the **residue** of f at z_0, where γ is a curve around z_0 with no other singularities in the interior. If f has a removable singularity, the residue is 0, and if it is a pole of order n, the residue may be computed as

$$\text{Res}(f, z_0) = \lim_{z \to z_0} \frac{1}{(n-1)!} \frac{d^{n-1}}{dz^{n-1}} (z - z_0)^n f(z).$$

Theorem 4.5.2 (Residue Theorem)**.** *Let f be meromorphic. Then, for any simple contour γ oriented counterclockwise,*

$$\int_\gamma f(z)dz = 2\pi \sum_a Res(f, a),$$

where a ranges over all singularities of f inside γ.

Definition 4.5.2. A holomorphic function f is said to be of **exponential order** n if n is the infimum of a such that $f(z) \ll e^{|z|^n}$.

Theorem 4.5.3 (Weierstrass/Hadamard Factorization Theorems)**.** *If f is an entire function with nonzero zeros a_n and a zero of order m at 0, there exists an entire function g and a sequence of integers n such that*

$$f(z) = z^m e^{g(z)} \prod_{n=1}^{\infty} E_{p_n}(\frac{z}{a_n}),$$

where

$$E_n(z) = (1-z)e^{\sum_{i=1}^{n} \frac{z^n}{n}},$$

where the convergence of the product is normal on $\mathbb{C}$. Moreover, if f is of order n, then it suffices to take $g(z)$ to be a polynomial of degree n and $p_k = p$, where p is the smallest integer such that

$$\sum_{n=1}^{\infty} \frac{1}{|a_n|^{p+1}}$$

converges.

Remark 4.5.1. Note that the convergence of the product is guaranteed since $|E_p(\frac{z}{a_n}) - 1| \ll |\frac{z}{a_n}|^{n+1}$.

Theorem 4.5.4 (Mittag-Leffler Theorem)**.** *Any meromorphic function $f : U \to \mathbb{C}$ with a set of singularities E without a limit point in U can be written as $f = g + h$, where g is holomorphic on U and for any $a \in E$, $h - p_a(z)$ has a removable singularity at a, where $p_a(z)$ is the principal part of h at a. In particular, every function has the normally convergent expansion*

$$f(z) = h(z) + \sum_{n=1}^{\infty} p_{a_n}(z) + g_n(z),$$

where $g_n(z)$ are polynomials chosen to make the sum converge. Particularly, if f only has simple poles, is defined at 0, and is uniformly bounded on a sequence of circles with radii tending to ∞, then

$$f(z) = f(0) + \sum_{n=1}^{\infty} \frac{b_n}{z - a_n} + \frac{b_n}{a_n},$$

where b_n is the residue of f at a_n.

Example 4.5.2. (a) $\sin z$ is a entire function of order 1 with a zero of order 1 at 0. Since the zeros are $\pm n\pi$, and thus are asymptotic to the harmonic series, clearly $p = 1$ suffices. Thus, by Hadamard,

$$\sin z = ze^{az+b} \prod_{n=1}^{\infty} (1 - \frac{z}{n\pi})e^{\frac{z}{n\pi}}(1 + \frac{z}{n\pi})e^{-\frac{z}{n\pi}} = ze^{az+b} \prod_{n=1}^{\infty} (1 - \frac{z^2}{n^2\pi^2}),$$

and using the fact that $\sin z$ is odd and dividing by z and substituting 0 yields $a = b = 0$, i.e.

$$\sin z = z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n\pi^2}\right).$$

(b) By the same exact logic,

$$\cos z = e^{az+b} \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{(n-\frac{1}{2})^2\pi^2}\right),$$

where plugging in 0 and using the fact that $\cos z$ is even yields

$$\cos z = \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{(n-\frac{1}{2})^2\pi^2}\right) = \prod_{n=0}^{\infty} \left(1 - \frac{4z^2}{(2n+1)^2\pi^2}\right).$$

(c) $e^z - e^{iz}$ is an entire function of order 1. It has zeros whenever $e^{z(1-i)} = 0$, i.e. $z = \frac{2\pi n}{1-i}$. Then,

$$e^z - e^{iz} = ze^{az+b} \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{(\frac{2\pi n}{1-i})^2}\right),$$

and by standard techniques we conclude that $b = \ln(1-i) = \sqrt{2} - \frac{\pi}{4}$ and since $e^{z-az} - e^{iz-az}$ is odd, $-z(1-a) = z(i-a)$, i.e. $a - 1 = i - a \implies a = 1 + i$, so

$$e^z - e^{iz} = ze^{(1+i)z+(1-i)} \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{(\frac{2\pi n}{1-i})^2}\right).$$

Example 4.5.3. (a) $\tan z$ has simple poles at $z = \pi(\pm n \pm \frac{1}{2})$ for $n \geqslant 0$ and satisfies the uniform boundedness condition, so by Mittag-Leffler, with residue -1 at every pole, one has

$$\tan z = \sum_{n=0}^{\infty} -\frac{1}{z - \pi(n+\frac{1}{2})} - \frac{1}{z - \pi(-n-\frac{1}{2})} = \sum_{n=0}^{\infty} \frac{8z}{\pi^2(2n+1)^2 - 4z^2}.$$

Alternatively, one may use the Hadamard factorization for $\cos z$ and the fact that $\tan z$ is the logarithmic derivative to conclude

$$-\tan z = \sum_{n=0}^{\infty} \frac{-8z}{(2n+1)^2\pi^2 - 4z^2},$$

which yields the same expansion.

(b) Doing term by term differentiation of the above series yields $\sec^2 z = \sum_{n=0}^{\infty} \frac{8(\pi^2(2n+1)^2+4z^2)}{(\pi^2(2n+1)^2-4z^2)^2}$.

4.6 Montel and Runge

We now define a version of compactness for families holomorphic functions.

Definition 4.6.1. A family $\mathcal{F} \subset H(U)$ is said to be **normal** if every sequence in $\mathcal{F}$ has a uniformly convergent subsequence.

Montel's theorem provides a simple description of normal families:

Theorem 4.6.1 (Montel's Theorem)**.** *A family $\mathcal{F} \subset H(U)$ is normal iff it is locally uniformly bounded.*

Proof. The forward direction is essentially an application of Arzela-Ascoli, using the bounds on the derivative $f'(z)$ from $f(z)$. The converse follows since precompact sets are bounded. □

An interesting parallel to Montel's Theorem is the following statement:

Theorem 4.6.2 (Fundamental Normality Test)**.** *A family $\mathcal{F} \subset H(U)$ of functions all omitting the same two values $a, b \in C$ is normal.*

Finally, we want to provide an analogue of such convergence for meromorphic functions, provided in the form of Runge's theorem:

Theorem 4.6.3 (Runge's Theorem)**.** *If $f \in H(U)$ and A is a set with at least one value from each connected component of $\mathbb{C}\backslash K$, where $K \subset U$ is compact, then there exists a sequence of rational functions with poles in A that converge uniformly to f on K.*

4.7 Automorphisms of Riemann Surfaces

Recall the construction of the Riemann sphere as $\mathbb{CP} = \mathbb{C} \cup \{\infty\}$. We are interested in studying the structure of simply-connected Riemann surfaces (complex manifolds). This is made extremely easy with the following theorem:

Theorem 4.7.1 (Uniformization Theorem)**.** *Every simply-connected Riemann surface is conformally equivalent to $\mathbb{CP}, \mathbb{C}$, or $\mathbb{D}$.*

Thus, it suffices to understand holomorphic/meromorphic maps and automorphism between each of these three Riemann surfaces. We first focus on $\mathbb{CP}$.

Definition 4.7.1. A **Möbius transformation** is a map of the form

$$f(z) = \frac{az+b}{cz+d},$$

satisfying the following properties:

(a) The Möbius transformations form a group isomorphic to $\mathbb{P}\mathrm{GL}_2(\mathbb{C})$ according to

$$\frac{a+bi}{c+di} \Longleftrightarrow \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

(b) A Möbius transformation is uniquely defined by three points.

(c) If $f(z) = \frac{az+b}{cz+d}, f^{-1}(z) = \frac{dz-b}{-cz+a}$

Proposition 4.7.1. *Let $f : \mathbb{CP} \to \mathbb{CP}$ be meromorphic. Then, f is a rational function.*

Proof. Since $\mathbb{CP}$ is compact, any nonconstant f has finitely many zeros. Moreover, f has finitely many singularities, and each singularity is a pole (otherwise f is not meromorphic on $\mathbb{C}$, i.e. holomorphic on $\mathbb{CP}$). Multiplying by the poles and dividing by the zeros yields a bounded function on $\mathbb{C}$, so by Liouville it is constant. Thus, f is rational. ☐

Corollary 4.7.1. *All diffeomorphisms (that is, complex automorphisms) of $\mathbb{CP}$ are Möbius transformations.*

Proof. We know that f has exactly one zero and one pole and is rational. ☐

Proposition 4.7.2. *The diffeomorphisms of $\mathbb{C}$ are linear functions.*

Proof. Note that by Casorati-Weierstrass, f does not have an essential singularity at ∞, since the image of f in a neighborhood of ∞ is not dense in $\mathbb{C}$ by the open mapping theorem. Thus, f is a polynomial. The only injective polynomials are linear functions. The result then follows. □

Proposition 4.7.3. *The automorphisms of* $\mathbb{D}$ *are the* ***Blaschke factors*** $f(z) = e^{i\theta}\frac{z-a}{1-\bar{a}z}$.

Proof. Suppose $f(0) = 0$. Applying the Schwarz lemma to f, f^{-1}, we get that $|f'| = 1$, so $f(z) = e^{i\theta}z$. Otherwise, map the zero a of f to 0 using a Blaschke product. □

Remark 4.7.1. Blaschke factors are very special because they replace a zero at $\frac{1}{\bar{a}}$ with a pole at a, and also have magnitude 1 on the unit circle, so multiplying by them does not change the magnitude of the unit circle.

Proposition 4.7.4. *There are no holomorphic maps* $\mathbb{CP} \to \mathbb{C}, \mathbb{C} \to \mathbb{D}$.

Proof. $\mathbb{CP}$ is compact and $\mathbb{C}$ is not. The latter is the statement of Liouville's theorem. □

4.8 Jensen's Formula and Bounds on Zeros

Since complex functions are so well-behaved, it is natural to ask if one may obtain certain bounds on their growth as it relates to the number of zeros. One first needs the following absolutely fundamental lemma.

Lemma 4.8.1 (Schwarz Lemma). *Let* $f : \mathbb{D} \to \mathbb{D}$ *be holomorphic and* $f(0) = 0$. *Then,* $|f(z)| \leqslant |z|$, *and* $|f'(0)| \leqslant 1$. *If equality holds in either case,* $f(z) = e^{i\theta}z$.

Proof. Define $g(z) = \frac{f(z)}{z}$. Then, $\lim_{|z|\to 1} |g(z)| = 1$, so by the maximum principle on $\{|z| = r\}$ and sending $r \to 1$, we can conclude that $|g(z)| \leqslant 1$, i.e. $|f(z)| \leqslant |z|$ and $|f'(0)| = |g(0)| \leqslant 1$. If equality holds, then by open mapping theorem/maximum principle, f has constant magnitude and so $f(z) = e^{i\theta}z$. □

Lemma 4.8.2 (Schwarz-Pick Lemma). *If* $f : \mathbb{D} \to \mathbb{D}$ *is holomorphic,*

$$\left|\frac{f(z_1) - f(z_2)}{1 - \overline{f(z_1)}f(z_2)}\right| \leqslant \left|\frac{z_1 - z_2}{1 - \overline{z_1}z_2}\right|,$$

with equality holding iff f *is a Blaschke factor, and*

$$\frac{|f'(z)|^2}{1 - |f(z)|^2} \leqslant \frac{1}{1 - |z|^2}.$$

Proof. Consider the Blaschke factors τ_2, τ_2 sending $0 \to z_1, 0 \to f(z_1)$, respectively. Then, $\tau_2 \circ f \circ \tau_1^{-1}(z)$ satisfies the assumptions of the Schwarz lemma, so replacing $z \to \tau(z_2)$ yields the desired inequality, and sending $z_1 \to z_2$ yields the latter inequality. □

Definition 4.8.1. The **pseudohyperbolic metric** on $\mathbb{D}$ is defined as $\rho(z, w) = \left|\frac{z-w}{1-\bar{z}w}\right|$. The Schwarz-Pick lemma implies that with respect to this metric, analytic functions on $\mathbb{D}$ are Lipschitz.

Theorem 4.8.1 (Borel-Caratheodory). *An entire function is bounded by its real part according to*

$$\sup_{|z|\leqslant r} |f| \leqslant \frac{2r}{R-r}M + \frac{R+r}{R-r}f(0)$$

for $M = \sup_{|z|\leqslant R} \operatorname{Re} f$.

Proof. If f is nonconstant, the idea is to apply the Schwarz lemma. Suppose $f(0) = 0$, and note that since $\operatorname{Re} f$ is a harmonic nonconstant function, $M > 0$. Then, the image of f lies in the shifted half-plane $\operatorname{Re} z \leqslant M$, which can be mapped to the disk of radius R using

$$\phi(z) = \frac{Rz}{z - 2M}.$$

Thus, Schwarz lemma yields that for $|z| \leqslant r$,

$$\left| \frac{Rf(z)}{f(z) - 2M} \right| \leqslant r \implies \sup_{|z| \leqslant r} |f| \leqslant \frac{2r}{R - r} M.$$

In the general case, just apply the proof to $f(z) - f(0)$. □

One of the most important functions is the complex logarithm, defined as $\log z = \log |z| + i\operatorname{Arg}(z)$. The issue is that the argument of a complex number is not well-defined up to multiples of 2π, requiring a **branch cut** where $\log z$ is undefined. The standard choice is to make the branch cut at the negative x-axis and let $\operatorname{Arg}(z) \in (-\pi, \pi)$, which corresponds to the Log z. However, one may shift the branch cut as is necessary, as long as the domain does not contain a curve around 0.

The key theorem that relates the zeros of a holomorphic function to its growth is known as Jensen's formula:

Theorem 4.8.2 (Jensen's formula)**.** *For a meromorphic function f on U and γ a circle of radius R centered at z_0 and contained in U,*

$$\log |f(z_0)| = \frac{1}{2\pi} \int_\gamma \log |f(z)| dz + \sum_k \log \frac{|a_k - z_0|}{R} - \sum_k \log \frac{|b_k - z_0|}{R},$$

where a_k and b_k are the zeros and poles of f in the interior of γ, respectively.

Proof. By scaling and shifting, one may assume that $\gamma = \partial \mathbb{D}$, i.e. $z_0 = 0, R = 1$. Multiplying f by appropriate Blaschke factors makes f nonzero holomorphic on $\mathbb{D}$ at no cost on the right, since the Blaschke factors have magnitude 1 on the boundary and $\log 1 = 0$, and a cost of $\log |\rho|$ on the left (since the Blaschke product is evaluated at 0). Thus, the problem reduces to showing

$$\frac{1}{2\pi} \int_{\partial \mathbb{D}} \log |f| dz = \log(|f(0)|),$$

which follows since $\log |f| = \operatorname{Re} \log f$ is a harmonic function and thus satisfies the mean value property. □

Corollary 4.8.1. *The most useful version of Jensen's formula is that for holomorphic functions, which states*

$$\frac{1}{2\pi} \int_\gamma \log |f(z)| dz = \log |f(z_0)| + \sum_k \log \frac{R}{|a_k - z_0|}.$$

Crucially, this formula relates the number of zeros to the growth of an entire function.

Proposition 4.8.1. *If f is entire of order A, if N is the number of zeros of f in $B(0, R)$, then $N \ll R^A$.*

Proof. Consider Jensen's formula applied to a circle of radius $2R$. Then,

$$\frac{1}{2\pi}\int (2R)^A dz \geqslant \log|f(0)| + \sum_k \log\frac{2R}{|a_k|}.$$

For every zero in $B(0,R)$, the term in the sum is at least $\log 2$, and for all other zeros the term is nonnegative, so

$$N\log 2 \ll R^A.$$

□

Example 4.8.1. There is no nonzero entire f such that $f \ll e^{|z|}$ and $f(n^{\frac{1}{3}}) = 0$ for all $n \geqslant 0$. This is because $N \gg R^3 \gg R^1$, and 1 is the order of f.

4.9 Phragmen-Lindelof

Often times, one wants to bound a holomorphic function defined on some unbounded region in the complex plane. For bounded regions, one may appeal to the maximum modulus principle, and for unbounded ones, the argument of the Hadamard three-lines lemma motivates the following set of propositions.

Proposition 4.9.1 (Phragmen-Lindelof)**.** *If $f \ll e^{e^{\frac{\pi}{b-a}|\operatorname{Im} z|}}$ is holomorphic in the strip $a < \operatorname{Re} z < b$, bounded by M on the edges of the strip, then it is bounded by M everywhere, i.e. f satisfies the maximum principle.*

Proof. Multiplying by $e^{-\epsilon e^z}$, using the maximum modulus principle, and sending $\epsilon \to 0$ concludes the proof. □

Corollary 4.9.1 (Phragmen-Lindelof for Sectors)**.** *If f is holomorphic in the sector $\alpha \leqslant \theta \leqslant \beta$ and of exponential type at most $\frac{\pi}{\alpha-\beta}$, then f satisfies the maximum principle.*

Proof. Apply Phragmen-Lindelof to $f(e^{iz})$. □

4.10 Injective, Proper Functions, and Blaschke Products

Proposition 4.10.1. *If $f : U \to V$ is holomorphic injective, then f' does not vanish on U.*

Proof. WLOG, suppose $z_0 = 0$. Then, write $f(z) = a_k z^k + ... = z^k h(z)$ for some analytic function h such that $h(0) \neq 0$. This implies that locally, h has a holomorphic k-th root, so $f(z) = (zg(z))^k$. But the image of $zg(z)$ contains a neighborhood of zero by the Open Mapping Theorem, so there exist two distinct angles θ_1, θ_2 such that $g(z_1) = re^{i\theta_1}, g(z_2) = re^{i\theta_2}$, and $z_1 g(z_1)^k = z_2 g(z_2)^k$ contradicting the injectivity of f. □

The Inverse Function Theorem guarantees that if $f : U \to V$ is holomorphic and bijective, then the derivative of f^{-1} satisfies C-R and so f is conformal.

To analyze the zeros of holormorphic functions, we have two very powerful tools:

Theorem 4.10.1 (Argument Principle)**.** *For f meromorphic in U and γ a simple closed curve, the value of $\frac{1}{2\pi i}\oint_\gamma \frac{f'(z)}{f(z)}dz$ = number of zeros - number of poles inside U.*

Theorem 4.10.2 (Rouche's Theorem)**.** *If $f, g \in H(U) \cap C(\overline{U})$ and and $|g| < |f|$ on ∂U, f and $f+g$ have the same number of zeros in U.*

Definition 4.10.1. A map $f : U \to V$ is **proper** if the preimage of any compact set in V is compact in U.

Theorem 4.10.3. *(The Fundamental Theorem of Blaschke Products)*

(a) A map $f : \mathbb{D} \to \mathbb{D}$ is proper iff it is a finite Blaschke product.

(b) Given a sequence $a_n \in \mathbb{D}$ such that $\sum_{n=1}^{\infty}(1 - |a_n|) < \infty$, there exists a function $f = \prod_{n=1}^{\infty} \frac{|a_n|}{a_n} \frac{z-a_n}{1-\overline{a_n}z} \in H(\mathbb{D})$ that vanishes precisely on $\{a_n\}$.

Proof. Suppose f is proper. Then, $f^{-1}(0)$ is finite, so f has finitely many zeros. Moreover, $f^{-1}(\overline{B(0,r)})$ for $r < 1$ is compact, and so avoids the boundary of the $\mathbb{D}$. Thus, $\lim_{|z|\to 1}|f(z)| = 1$. Divide by the Blaschke factors corresponding to those zeros to obtain a map $\widetilde{f}$ that does not vanish on the unit disk and extends to a function of constant modulus on the boundary. The image of a function of constant modulus is a subset of a circle, so by the Open Mapping Theorem, the function is constant. Thus,

$$f = e^{i\theta} \prod_{i=0}^{n} B_i,$$

where B_i are the Blaschke factors of f. Conversely, if f is of the above form, for any compact set $K \subset \overline{B(0,r)}, r < 1$, $f^{-1}(K)$ avoids the boundary, and so is closed and bounded, i.e. compact. Finally, for a function $f \in H(\mathbb{D})$ with zeros $\{a_n\}$, define the partial products as above. The ☐

4.11 Harmonic Functions and Laplace's Equation

Definition 4.11.1. Let $U \subset \mathbb{R}^n$ be open. Then, $f : U \to \mathbb{C}$ is **harmonic** if $\Delta u = 0$ on U.

Lemma 4.11.1. *If f is holomorphic on U, then $\mathrm{Re}(f), \mathrm{Im}(f)$ are harmonic on U. Conversely, on a simply connected subset of $\mathbb{C}$, every harmonic function is the real/complex part of a holomorphic function, unique up to a constant.*

Proof. One direction follows easily from C-R. Conversely, for u harmonic, define $f = u_x - iu_y$. Then, f has a primitive $F(z) = f(z_0) + \int_{z_0}^{z} f(\zeta)d\zeta$ for some $z_0 \in U$. If $U(z) = \mathrm{Re}\, F$, $F'(z) = U_x - iU_y = u_x - iu_y$, so $U(z) = u(z)$ for all z. ☐

Remark 4.11.1. This does not necessarily hold for non-simply connected regions. As a counterexample, $\log\sqrt{x^2+y^2}$ is harmonic in the punctured plane but is not a real part of an analytic function. If it did, then $\frac{x}{\sqrt{x^2+y^2}} - i\frac{y}{\sqrt{x^2+y^2}}$ would have a primitive. But this function is not path-independent and therefore not conservative.

Corollary 4.11.1. *If f is nonzero holomorphic on a simply connected open U, $\log|f|$ exists and is harmonic on U.*

Proof. Take the real part of the antiderivative of $\frac{f'}{f}$. ☐

Theorem 4.11.1 (Analytic Continuation)**.** *If $f : U \to \mathbb{C}$ is holomorphic and $U \subset V$, then, there exists an analytic continuation of f*

Harmonic functions enjoy most of the same properties as holomorphic functions.

Theorem 4.11.2. *(a) A function $f : U \to \mathbb{R}$ is harmonic iff for any ball $B(a, r) \subset U$,*

$$\frac{1}{\mu(B(a,r))}\int_{B(a,r)} f(x)dx = \frac{1}{\mu(\partial B(a,r))}\int_{\partial B(a,r)} f(x)dx = f(a).$$

(b) A harmonic function on $\mathbb{R}^n$ bounded above or below is constant.

(c) (Strong Maximum Principle) If U is connected and $f : U \to \mathbb{R}$ achieves a local maximum or minimum, then f is constant.

(d) (Weak Maximum Principle) If U is bounded, connected, and f is harmonic and continuous up to $\overline{U}$, f achieves its maximum and minimum on ∂U.

(e) (Identity Theorem) Two harmonic functions $f : U \to \mathbb{R}$ that agree on $V \subset U$ open agree on U.

(f) A harmonic function is smooth.

Proof. Note that it suffices to prove the mean value property for spheres. WLOG, suppose $a = 0$. Then, by the divergence theorem, we compute

$$\frac{d}{dr}\left[\frac{1}{\mu(\partial B(0,r))}\int_{\partial B(0,r)} f(x)dx\right] = \frac{1}{\mu(B(0,1))}\int_{\partial B(0,1)} \nabla f(rx)\cdot x dx = \frac{1}{\mu(B(0,1))}\int_{B(0,1)} \Delta f(rx)dx = 0$$

if f is harmonic. Conversely, the mean value property implies that Δf vanishes on arbitrarily small balls, and so $\Delta f = 0$. Since at $r = 0$, this function approaches $f(a)$, the claim follows.
WLOG suppose a harmonic function is nonnegative. Then, for $x, y \in U$, pick R_1, R_2 such that $R_2 = R_1 + 2|x - y|$, i.e. so that $B(x, r_1) \subset B(y, r_2)$. Then,

$$f(x) = \frac{1}{\mu(B(x,R_1))}\int_{B(x,R_1)} f(t)dt \leqslant \frac{\mu(B(y,R_2))}{\mu(B(x,R_1))}\frac{1}{\mu(B(y,R_2))}\int_{B(y,R_2)} f(t)dt = f(y),$$

and as $R_1 \to \infty$, the ratio of the volumes tends to 1, i.e. $f(x) \leqslant f(y)$. By symmetry, f is constant. The identity principle can be proven by showing that the set of points where an analytic function locally vanishes is both closed and open. If U is connected and f achieves its local maximum/minimum inside U, then f is locally constant, so by the identity principle, it is globally constant. □

Remark 4.11.2. Note that the mean value formula immediately implies that a normally convergent sequence of harmonic functions is harmonic, and by the monotone convergence theorem, that a decreasing sequence of harmonic functions is harmonic.

Remark 4.11.3. Note that the maximum of two harmonic functions is **not necessarily harmonic.**

4.12 Subharmonic Functions

If we relax the equality in the mean value formula to an inequality, we obtain so-called subharmonic functions.

Definition 4.12.1. TFAE for an upper semi-continuous $f : U \to \mathbb{R} \cup \{-\infty\}$:

(a) For all $B(a, r) \subset U$,

$$f(a) \leqslant \frac{1}{\mu(B(a,r))}\int_{B(a,r)} f(x)dx.$$

(b) For all $B(a,r) \subset U$,

$$f(a) \leqslant \frac{1}{\mu(\partial B(a,r))} \int_{\partial B(a,r)} f(x)dx.$$

(c) If U is a bounded open set, for every harmonic h on $V \subset U$ continuous up to the boundary one has $f|_{\partial V} \leqslant h|_{\partial V}$, then $f \leqslant h$ in V.

(d) If f is C^2, $\Delta f \geqslant 0$ in U.

If any of these hold, f is called **subharmonic.** The negative of a subharmonic function is called a **superharmonic function.**

Proof. For (b) $\implies$ (c), suppose $f|_{\partial V} \leqslant h|_{\partial V}$ but $f(a) > h(a)$ for some harmonic function h and $a \in B$. Then, the set where $f - h$ is positive is open and nonempty. Suppose x is the maximum of $f - h$. Then, the sub-mean value property implies that $f - h$ is constant in a neighborhood of x, implying that the set where $f - h$ achieves its maximum is open and closed, i.e. $f - h$ is constant, which is a contradiction. For (c) $\implies$ (b), take a harmonic function h such that $h|_{\partial B} = f|_{\partial B}$ (which can be done by Poisson's formula). Then, $f(a) \leqslant h(a) = \frac{1}{\mu(\partial B(a,r))} \int_{\partial B(a,r)} f(x)dx$.

(b) $\implies$ (a) follows by integrating on both sides, and (a) $\implies$ (b) by continuity.

Finally, if f is C^2, the argument in the properties of harmonic functions directly proves the equivalence of (a) and (c). □

Remark 4.12.1. The set $\{x : f(x) = -\infty\}$ for a subharmonic function f has measure zero. This follows from the following facts:

(a) If f is subharmonic, $f \in L^1_{loc}$, since the set of points where f is locally integrable is both open and closed.

(b) If f is subharmonic, $\int_{\partial B} f dx > -\infty$ if $B \subset U$. This is true since the integral over a sphere of f can be bounded by the value of f at any point of inside the ball, which implies the integral is finite, for otherwise $f = -\infty$ on an open set.

Corollary 4.12.1. *From the third definition, one obtains the maximum principle for subharmonic functions: if $f : U \to \mathbb{R}$ is subharmonic and achieves a* ***global maximum*** *in U, it is constant. Moreover, if U is bounded and f is continuous up to ∂U, f achieves its maximum on ∂U. As a counterexample to the local maximum principle,* $\max(x, 0)$ *is subharmonic, yet has local maxima in the left half-plane.*

Proposition 4.12.1. *(a) Subharmonic functions form a positive cone, i.e. if u, v are subharmonic, $a, b \geqslant 0$, then $au + bv$ is subharmonic.*

(b) If $u_1, ..., u_n$ are subharmonic, $\max(u_1, ..., u_n)$ *is subharmonic.*

(c) If ϕ is convex harmonic and u is subharmonic, then $\phi \circ u$ is subharmonic. In particular $e^u, u^+ = \max(u, 0)$, and u^p, $p \geqslant 0$, are subharmonic functions.

(d) Since it suffices to check that $\Delta u \geqslant 0$ to show u is subharmonic, $\log(1 + |f|^2)$ *is subharmonic for f holomorphic.*

(e) From the sub-mean value property, it is clear that a normally convergent sequence of subharmonic functions is subharmonic, and by monotone convergence, a decreasing sequence of subharmonic functions is subharmonic.

Note that if f is defined on a simply connected domain (possibly with zeros), then $\log|f|$ is subharmonic if we define $\log 0 = -\infty$. This is because subharmonicity is a local property, i.e. being subharmonic in a neighborhood of every point implies global subharmonicity.
Here is an important analogue of Liouville's theorem, which now cruically holds only in $\mathbb{R}^2$.

Proposition 4.12.2. *A subharmonic function u on $\mathbb{C}$ that is bounded above is constant.*

Proof. Consider a perturbation of u defined by $u_\epsilon(z) = u(z) - \epsilon \log|z|$. This perturbation agrees with u on $\partial\mathbb{D}$, and $u(z) \leqslant u_\epsilon(z)$ on $|z| > 1$. Moreover, by construction,

$$\sup_{|z|>1} |u_\epsilon| \leqslant \sup_{|z|=1} |u_\epsilon| = \sup_{|z|=1} |u| = \sup_{\mathbb{D}} |u|,$$

where we can use the maximum principle on u_ϵ since it goes to $-\infty$ as $|z| \to \infty$, so

$$u(z) = u_\epsilon(z) + \epsilon \log|z| \leqslant \sup_{\mathbb{D}} |u| + \epsilon \log|z|$$

on $|z| > 1$. Sending $\epsilon \to 0$ yields $u(z) \leqslant \sup_{\mathbb{D}} |u|$ on $\mathbb{C}$, which violates the maximum principle, a contradiction. Thus, u is constant. □

Subharmonic functions enjoy very nice properties when discussing their means.

Lemma 4.12.1. *A radial function f is subharmonic iff f is a convex increasing function of* $\log r$.

Proposition 4.12.3. *Let $u : \mathbb{C} \to \mathbb{R}$ be subharmonic and and define $I_u(r), J_u(r), M_u(r)$ to be the spherical mean, ball mean, and maximum value of u for $B(0, R)$ respectively. Then, I_u, J_u, M_u are convex increasing continuous functions of* $\log r$, *and $u(0) \leqslant J_u(r) \leqslant I_u(r)$ and $u(0) = J_u(0) = I_u(0) = M(0)$.*

Remark 4.12.2. Note that I_u, J_u are well-defined since subharmonic functions are locally integrable by Remark 6.9.

Proof. We first prove the statement for M_u. We use the following characterization of convex functions $f : \mathbb{R} \to \mathbb{R}$ - f is convex iff for any linear function l, $f - l$ attains its maximum on the boundary. Then, for any $a, b \in \mathbb{R}$, note that

$$v(z) := u(z) - a \log|z| - b$$

is subharmonic in an annulus around 0 (since $\log|z|$ is harmonic), and so by the maximum principle, if $M_u(r) - a \log r - b \leqslant 0$ on the boundary, then $v \leqslant 0$ on the boundary, and therefore also on the annulus, and since $M_u(r) = \sup_{|z|=r} v(z) - a \log r - b$, we conclude that $M_u(r) - a \log|z| - b \leqslant 0$ on the annulus, showing that M_u is convex as a function of $\log r$.

To prove the claim for I_u, J_u, we first assume $u \in C^2$. Then, since $\Delta = \partial_{rr} + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_{\theta\theta}$, and I_u is radial, we have that $\int \Delta u = \Delta \int u \geqslant 0$, which implies

$$I_u''(r) + \frac{1}{r} I_u'(r) \geqslant 0 \iff (r\partial_r)^2 I_u(r) \geqslant 0 \iff t \to I_u(e^t) \text{ is convex,}$$

since $r = e^t$ implies $r\partial_r = \partial_t$. In the general case, we construct

$$u_\epsilon := \int I(x - \delta z)\phi(z)dz,$$

where ϕ is a smooth nonnegative radial function with $\int \phi = 1$ that equals 1 on $\partial\mathbb{D}$. Then, u_ϵ decreases to u by the sub-mean value property, so $I_\epsilon := I_{u_\epsilon}$ decreases to I, so monotone convergence implies convexity of I_u in $\log r$ for arbitrary subharmonic u. One can also easily see that I_u is increasing as follows: find a monotone sequence of continuous functions g_k decreasing to u. Then, for $r_1 < r_2$, pick h harmonic so that $u \leqslant h = g_k$ on $|z| = r_2$, then $u \leqslant h$ everywhere and $I_u(r_1) \leqslant I_h(r_1) = I_h(r_2) = I_{g_k}(r_2)$. Sending g_k to infinity and using monotone convergence completes the proof, and the same argument applies for J_u. The fact that $J_u(r) \leqslant I_u(r)$ follows from the fact that J_u is obtained by radially integrating I_u, which is increasing, and continuity for all functions follows from convexity. □

Corollary 4.12.2. *If $f \in H(\mathbb{D})$, then $I_{\log|f|}, I_{\log^+|f|}(e^t)$ are increasing continuous function convex in t, so $I_{\log|f|}, I_{\log^+|f|} \to \infty$.*

Corollary 4.12.3. *If u is harmonic in an annulus, then $I_u(r) = J_u(r) = a\log r + b$ since $\pm I_u$ are convex in $\log r$.*

Corollary 4.12.4. *Note that by setting $u = \log|f|$, the convexity of M_u directly implies the Hadamard three-circles lemma.*

Remark 4.12.3. The same exact argument shows that the radial L^p averages $I_{u,p} := \frac{1}{2\pi}\left(\int_0^{2\pi} u(re^{i\theta})^p d\theta\right)^{\frac{1}{p}}$ are increasing continuous functions convex in $\log r$.

We are now ready to state our main result.

Theorem 4.12.1 (Fundamental Theorem of Subharmonic Functions on $\mathbb{C}$)**.** *Let $u : \mathbb{C} \to \mathbb{R}$ be subharmonic. Then, if $\liminf_{r\to\infty} \frac{M_u(r)}{\log r} = 0$, then u is constant.*

Proof. This follows immediately from the following lemma.

Lemma 4.12.2. *If $\liminf_{x\to\infty} \frac{f(x)}{x} = 0$ and f is convex and increasing, then f is constant.*

Proof. By convexity, for $x < y < z$,

$$f(y) \leqslant \frac{z-y}{z-x} f(x) + \frac{(y-x)z}{z-x}\frac{f(z)}{z}.$$

Taking a subsequence such that $\frac{f(z)}{z} \to 0$ as $z \to \infty$ yields $f(y) \leqslant f(x)$, so f is decreasing and therefore constant. □

Since $M_u(\log r) \leqslant M_u(r)$ for large r, the conclusion then immediately follows. □

Remark 4.12.4. In this proof, you have to be careful to ensure that $r \to \infty$ to say $\log r \ll r$.

4.12.1 Exercises

Problem 4.12.1. Suppose $f \in H(\mathbb{D})$ and $f(0) \neq 0$. Then, if $\inf_{|z|=r} |f(z)| > 0$, then $\frac{1}{2\pi}\int \log|f(re^{i\theta})| d\theta \geqslant \log|f(0)|$, and moreover, if f is continuous up to the boundary, the f cannot vanish on ∂D on a set of positive measure.

Proof. One can prove the first fact in two ways - either using Jensen's formula (which actually shows that $I_{\log|f|}(r)$ is linear in $\log r$) and the fact that $\sum_{a_k \in B(0,r)} \log\frac{|a_k|}{r} < 0$, or using the subharmonicity of $\log|f|$. The second fact follows from the fact that $I_{\log|f|}$ is increasing, so by Fatou's lemma,

$$-\infty < \limsup_{r\to 1} \int_{|z|=r} \log|f(re^{i\theta})|d\theta \leqslant \int_{|z|=1} \limsup_{r\to 1} \log|f(re^{i\theta})|d\theta = \int_{|z|=1} f(re^{i\theta})d\theta.$$

□

4.13 Poisson Kernel and Conformal Mappings

An important and absolutely fundamental question in PDE is that of solving Laplace's equation, i.e. finding a harmonic function u such that

$$\begin{cases} \Delta u = 0 \text{ in } U, \\ u = g \text{ on } \partial U, \end{cases}$$

for some open $U \subset \mathbb{R}^2$ and real function g. How can we utilize complex analysis techniques to solve this? First, we consider $U = B(0,1)$ to take advantage of symmetry. If $f(z) = \sum_{n=0}^{\infty} a_n z^n$ is holomorphic in U, since every harmonic function is the average of a holomorphic function and its conjugate, making the substitution $z = re^{i\theta}$ yields

$$u(z) = \frac{1}{2}(f(z) + \overline{f(z)}) = \frac{1}{2}\left(\sum_{n=0}^{\infty} a_n r^n e^{in\theta} + \sum_{n=0}^{\infty} \overline{a_n} r^n e^{-in\theta}\right).$$

Setting $r = 1$, we may solve for the coefficients by setting $a_n = \widehat{g}(n)$. Now, the partial sums are harmonic, and converge normally to g whenever, for example, g is continuous. This gives an explicit solution

$$u(z) = \frac{1}{2}(f(z) + \overline{f(z)}) = \frac{1}{2}\left(\sum_{n=0}^{\infty} \widehat{g}(n) r^n e^{in\theta} + \sum_{n=0}^{\infty} \overline{\widehat{g}(n)} r^n e^{-in\theta}\right),$$

which can be rewritten in terms of the **Poisson kernel**

$$P_r(\theta) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{in\theta} = \frac{1}{1 - re^{i\theta}} + \frac{1}{1 - re^{-i\theta}} - 1 = \frac{1 - r^2}{1 - 2r\cos\theta + r^2} = \operatorname{Re} \frac{1 + re^{i\theta}}{1 - re^{-i\theta}}$$

as

$$\begin{aligned} u(re^{i\theta}) &= \frac{1}{4\pi} \sum_{n=-\infty}^{\infty} r^{|n|} \int_{-\pi}^{\pi} g(t)(e^{in\theta - int\theta} + e^{-in\theta + int\theta})dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t) \sum_{n=-\infty}^{\infty} r^{|n|} e^{in(\theta - t)} dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\theta - t) g(e^{it})dt. \end{aligned}$$

Ok, but what about $\mathbb{R}^n$? There, the Poisson kernel is

$$P_r(x, \zeta) = \frac{r^2 - |x|^2}{r|x - \zeta|^n}$$

for $\zeta \in S^{n-1}$. Then, the Poisson integral formula becomes

$$u(x) = \frac{1}{\omega_{n-1}} \int_{S^{n-1}} P_r(x, \zeta) g(\zeta) d\zeta,$$

where $\omega_{n-1} = \mu(S^{n-1})$. What are the regularity conditions on g that allow you to use the Poisson integral formula?

Proposition 4.13.1. *If $g \in C(\partial\mathbb{D})$, then the $P_r[g] \to g$ uniformly as $r \to 1$. If $f \in L^1(\partial\mathbb{D})$, then the Poisson kernel $P[g]$ is harmonic in $\mathbb{D}$. Moreover, if $g \in L^p(\partial\mathbb{D})$, then $\|P_r[g]\|_p \leqslant \|g\|_p$ and $P_r[g] \to g$ in L^p.*

Proof. That the integral formula is harmonic follows directly from Morera's theorem and the fact that holomorphic functions are harmonic. By the maximum principle, $\|P[g]\|_\infty = \|g\|_\infty$, so approximating g by trigonometric polynomials g_k on the disk yields that $P[g_k] \to P$ uniformly. Note that this implies that $P_r[g] \to g$ uniformly as $r \to 1$. Using Jensen's and approximating by continuous functions then yields that $P_r[g] \to g$ in L^p. □

These results yield the following theorem:

Theorem 4.13.1. *If u is harmonic in $\mathbb{D}$ and $\sup_r \|u_r\|_p < \infty$, then if $p = 1$, $u|_{\partial\mathbb{D}}$ is a complex Borel measure, and for $1 < p$, $u|_{\partial\mathbb{D}} \in L^p$.*

Consequently, $g \in L^p$ iff $P[g]$ is harmonic with radial norms uniformly bounded in L^p (except for $p = 1$, when g might be a measure).

Theorem 4.13.2 (Harnack's Inequality)**.** *If f is harmonic on $B(0,1)$ and continuous up to a boundary, then*

$$\frac{1-r}{(1+r)^{n-1}} f(0) \leqslant f(x) \leqslant \frac{1+r}{(1-r)^{n-1}} f(0)$$

on $\partial B(0,r) \subset B(0,1)$. More generally,

$$\sup_\Omega f \ll_\Omega \inf_\Omega f,$$

independent of f.

Proof. Using Poisson's formula and the fact that $1 - r \leqslant |x - \xi| \leqslant 1 + r$ (since $x \in B(0,r)$), the kernel satisfies

$$\frac{1-r}{(1+r)^{n-1}} \leqslant \frac{1-r^2}{|x-\xi|^n} \leqslant \frac{1+r}{(1-r)^{n-1}},$$

and the rest follows from the mean value property. □

We have obtained an explicit solution to the Laplace equation on the unit disk. But what about arbitrary domains? There, one has to use conformal mappings.

Definition 4.13.1. A **conformal map** is a biholomorphic bijective map between two regions.

The existence of such maps is a fundamental result of complex analysis:

Theorem 4.13.3 (Riemann Mapping Theorem)**.** *Every simply connected open proper subset U of $\mathbb{C}$ is conformally equivalent to the open unit disk, with a unique map $f : U \to \mathbb{D}$ such that $f(z_0) = 0, f'(z_0) > 0$.*

Proof. We first need a quick lemma:

Lemma 4.13.1 (Hurwitz's Theorem)**.** *If $f_n \in H(U)$ is a sequence of injective functions converging normally to a nonconstant f, then f is injective.*

Proof. Suppose f is not injective. Then, $f - a$ has at least two zeros in U for some $a \in \mathbb{C}$. Find a curve γ encompassing at least the two zeros and avoiding any other zeros. Then, by the argument principle,

$$1 \geqslant \frac{1}{2\pi i} \oint_{\gamma} \frac{f_n'}{f_n - a} dz \to \frac{1}{2\pi i} \oint_{\gamma} \frac{f'}{f - a} dz \geqslant 2.$$

□

Now, for an arbitrary simply connected open proper U, consider the family $\mathcal{F} \subset H(U, \mathbb{D})$. For $a \notin U$, note that $\log(z - a) \in \mathcal{F}$ exists and is injective. Moreover, note that $\log(z) - \log(z_0) - 2\pi i$ is bounded away from 0 by continuity of $\log$. Now, consider

$$f(z) = \frac{1}{\log(z) - \log(z_0) - 2\pi i},$$

which is a bounded injective holomorphic function. After scaling and applying a unit disk transformation, one may assume that $f : U \to \mathbb{D}$. By Montel and Hurwitz, one may take the supremum of $|f'(z_0)|$ over $f \in \mathcal{F}$, which is still a bounded and surjective holomorphic function. Now, suppose F is not surjective and misses some $\alpha \in \mathbb{D}$. If ϕ_α is the corresponding disk automorphism, $G = \sqrt{\phi_\alpha \circ F}$ is injective, $G(z_0) = 0$, and $|G'(z_0)| > |F'(z_0)|$, since by an application of the Schwarz lemma to $\Phi = \phi_\alpha^{-1} \circ z^2 \circ \phi_{g(z_0)}^{-1}$, one gets $F'(z_0) = \Phi'(0)G'(z_0)$ and $|\Phi'(0)| < 1$.

□

Finally, we need a lemma regarding the preservation of the harmonic properties of functions:

Lemma 4.13.2. *If f is a (sub)harmonic function on U and $g : U \to V$ is a conformal map, then $f \circ g$ is (sub)harmonic on V.*

Proof.

$$\frac{1}{4}\Delta(f \circ g) = \partial_z \partial_{\bar{z}}(f \circ g) = (\Delta f \circ g)|g'|^2,$$

and the conclusion follows from the Laplacian characterization of (sub)harmonic functions. □

Thus, to solve Laplace's equation on an arbitrary domain, one just needs to first map it conformally to the unit disk, solve the Dirichlet problem on the disk, and map it back onto the desired domain. Here is a list of commonly used conformal maps:

(a) **Upper Half-Plane (Second Quadrant) to Unit Disk** $z \to \frac{z-i}{z+i}$.

(b) **Right Half-Plane to Unit Disk** $z \to \frac{z-1}{z+1}$.

(c) **Horizontal Strip** $0 < \operatorname{Im} z < \pi$ **to Upper Half-Plane** $z \to e^z$.

(d) **Quarter-Plane to Half-Plane** $z \to z^2$.

(e) **Rotation by** θ **degrees** $z \to e^{i\theta} z$.

(f) **Unit disk to complement of unit disk** $z \to \frac{1}{z}$.

Remark 4.13.1. The inverses of these maps give the reverse conformal maps.

Remark 4.13.2. To show that a region is mapped to another region, it is sufficient to show that the boundaries and one interior point are mapped to each other.

Remark 4.13.3. Note that since Möbius transformations are automorphisms of the Riemann sphere, they are conformal maps of regions in the complex plane where they are defined.

Another common type of qual problem is to evaluate a particular contour integral. Here are some general guidelines on which contours one should use:

(a) Integrals of the form
$$\int_{\mathbb{R}} x^\alpha \frac{P(x)}{Q(x)}dx, \int \log x \frac{P(x)}{Q(x)}dx$$
for P, Q polynomials and $|\alpha| < 1$ can be evaluated with a keyhole contour with a branch of logarithm defined away from the positive real axis. Note that at the bottom edge of the contour, one has to use $z = e^{2\pi i}t$ to get a factor of $e^{2\pi i\alpha}$. Additionally, **note that one has to use the appropriate choice for residues based on the branch cut - for instance, a pole at -1 has to take the form $e^{\frac{3\pi i}{2}}$, not $e^{-\frac{\pi i}{2}}$.**

(b) Trigonometric integrals of the form
$$\int_{-\pi}^{\pi} \frac{P(\sin\theta, \cos\theta)}{Q(\sin\theta, \cos\theta)}$$
for P, Q polynomials may be evaluated by making the substitution $\cos\theta = \frac{1}{2}(z + \frac{1}{z}), \sin\theta = \frac{1}{2i}(z - \frac{1}{z}), z = e^{i\theta}$ and using residue theorem.

(c) Integrals of the form
$$\int_{\mathbb{R}} \sin x \frac{P(x)}{Q(x)}dx, \int_{\mathbb{R}} \cos x \frac{P(x)}{Q(x)}dx$$
for P, Q polynomials may be evaluated by taking them as the imaginary (real) part of a complex integral.

Theorem 4.13.4 (Sokhotski-Plemelj Formula)**.** *If f is holomorphic, then*
$$\lim_{\epsilon\to 0} \int_{\mathbb{R}} \frac{f(y)}{(x_0 \pm i\epsilon) - y}dy - \int_{|x_0 - y| \geqslant \epsilon} \frac{f(y)}{x_0 - y}dy = \mp\pi i f(x_0).$$

Proof. Consider rectangular contours around x_0 that goes to $\operatorname{Im} z = \pm\epsilon$. Then, by Cauchy's integral formula,
$$\int_{|x_0-y|\geqslant\epsilon} \frac{f(y)}{x_0 - y}dy - \int_{|x_0-y|\geqslant\epsilon} \frac{f(y \pm i\epsilon)}{x_0 - (y \pm i\epsilon)}dy \mp i\pi \int_0^1 f(\mp\epsilon e^{i\pi\theta})d\theta = \mp 2\pi i f(x_0),$$
where γ is a semicircular arc around x_0 of radius ϵ. As $\epsilon \to 0$, the value on the semicircular arc approaches $\mp f(x_0)$. By the partial Fourier transform property, the second term on the left tends as $\epsilon \to 0$ to
$$\int_{\mathbb{R}} \frac{f(y \pm i\epsilon)}{x_0 - (y \pm i\epsilon)}dy = \int_{\mathbb{R}} \widehat{f}(\xi)e^{2\pi i x_0 \xi}d\xi = \int_{\mathbb{R}} \widehat{f(x \pm i\epsilon)}e^{2\pi i(x_0 \pm i\epsilon)\xi}d\xi = \int_{\mathbb{R}} \frac{f_\epsilon(y)}{(x_0 \pm i\epsilon) - y}dy,$$
where $f_\epsilon(z) = f(z + i\epsilon)$. Thus,
$$\int_{\mathbb{R}} \frac{f_\epsilon(y)}{(x_0 \pm i\epsilon) - y}dy - \int_{|x_0-y|\geqslant\epsilon} \frac{f(y)}{x_0 - y}dy = \mp\pi i f(x_0),$$
and since $f_\epsilon \to f$ uniformly as $\epsilon \to 0$, this completes the proof. □

4.14 Hardy Spaces and Nevanlinna Class

Definition 4.14.1. For $f \in H(\mathbb{D})$, define the norms $\|f\|_r := \left(\frac{1}{2\pi}\int_0^{2\pi}|f(re^{i\theta})|^p d\theta\right)^{\frac{1}{p}}$. Then, since $|f|^p = e^{p\log|f|}$ is a subharmonic function, $\|f\|_r$ is increasing in r and convex in $\log r$.

Definition 4.14.2. Define the **Nevalinna class** N of all $f \in H(\mathbb{D})$ such that

$$\lim_{r\to 1}\frac{1}{2\pi}\int_0^{2\pi}\log^+|f(re^{i\theta})|d\theta < \infty.$$

Lemma 4.14.1. *If $f \in N$, then the zeros of f satisfy the Blaschke condition $\sum_n 1-|a_n| < \infty$.*

Proof. WLOG suppose $f(0) \neq 0$. By Jensen's, $f \in N$ implies

$$|f(0)|\prod_n \frac{r}{|a_n|} \leqslant C,$$

so sending $r \to 1$ yields $\prod_n |a_n| \geqslant |f(0)|^{-1}C > 0$, so $\sum_n(1-|a_n|) < \infty$. □

Definition 4.14.3. The **Hardy space** H^p for $p > 0$ is the subspace of $f \in H(\mathbb{D})$ such that

$$\|f\|_{H^p} := \lim_{r\to 1}\|f\|_r < \infty,$$

with H^∞ being the space of bounded holomorphic functions on the unit disk.

Remark 4.14.1. One can show that if $p \geqslant 1$ and f_n is Cauchy in H^p, using Cauchy's integral formula that it converges locally uniformly, i.e. H^p is a Banach space.

Remark 4.14.2. One easily sees that $H^\infty \subset H^p \subset H^q \subset N$ for $0 < q \leqslant p$.

Lemma 4.14.2. $H^p \subset H(\mathbb{D}) \cap L^p$.

Proof.

$$\|f\|_p^p = \int_0^1\int_0^{2\pi}|f(re^{i\theta})|^p r dr d\theta \leqslant 2\pi\int_0^1 r\|f\|_{H^p}^p dr = \|f\|_{H^p}^p.$$

□

Proposition 4.14.1. *If B is the infinite Blaschke product corresponding to the zeros of $f \in H^p$, then $f = Bg$ for $g \in H^p$.*

Proof. B is well-defined since the zeros of $f \in H^p$ satisfy the Blaschke condition, and the partial products converge monotonically on $\mathbb{D}$, so one concludes by monotone convergence. □

Lemma 4.14.3. *For $f \in H^p$, $\lim_{r\to 1} f(re^{i\theta}) = f(e^{i\theta})$ is well-defined a.e. and in $L^p(\partial\mathbb{D})$. Moreover $f(re^{i\theta}) \to f$ in L^p.*

Proof. Define $\phi(g) = \int_{\partial\mathbb{D}} gf_r$ on L^q for $f_r(x) = f(rx)$. Then, by Banach-Alaouglu, one can showing that there is a weak$-*$ convergent subsequence $f_{r_j} \to f \in L^p$, which can be then be shown to converge pointwise a.e. Moreover, $f_{r_j} \leqslant Hf$ (the maximal function of f), and Hf is bounded in L^p, so by dominated convergence theorem, $f_r \to f$ in L^p. □

Definition 4.14.4. Define $H^p(\partial D) := \{f \in L^p(\partial D) : \widehat{f}(n) = 0, n < 0\}$.

Lemma 4.14.4. *The mapping $H^p(\mathbb{D}) \to H^p(\partial\mathbb{D})$ given by $f \to \lim_{r\to 1} f(re^{i\theta})$ is an isomorphism of Banach spaces.*

Proof. By properties of the Fourier transform, it is easy to see that $H^p(\partial D)$ is a closed subspace of $L^p(\partial D)$, therefore it is also a Banach space. One notes that $\lim_{r\to 1} f(re^{i\theta}) = \sum_{n\geqslant 0} a_n e^{in\theta}$, so this is indeed a well-defined map. Moreover, we have already shown that this map is an isometry, so it is injective and continuous. Finally, one can use the Poisson kernel to show that it is surjective, and conclude by the open mapping theorem. □

A List of Symbols

$\mathbb{R}$	Real numbers
$\mathbb{C}$	Complex numbers
$\mathbb{N}$	Natural numbers
$\mathbb{Z}$	Integers
$\mathbb{Q}$	Rational numbers
$\mathbb{K}$	The field of real or complex numbers
$\operatorname{Re} z$	Real part of a complex number z
$\operatorname{Im} z$	Imaginary part of a complex number z
$\overline{z}$	Complex conjugate of a complex number z
$B(a, r)$	An open ball centered at a of radius r
$\int f d\mu$	Lebesgue integral of f with respect to a measure μ
$\oint$	Contour integral over a closed contour
$\mathcal{F}\{f\}, \widehat{f}$	Fourier transform of f
$\mathcal{F}^{-1}\{f\}, \check{f}$	Inverse Fourier transform of f
∂A	Boundary of a set A
sign	Signum function
X^*, T^*	Continuous dual of a Banach space X/Adjoint of an operator T
$\mathcal{B}(X, Y)$	Space of bounded linear operators $T : X \to Y$
$\langle \xi \rangle$	Japanese bracket of ξ
H^p	Sobolev space, Hardy space
Δ	Laplacian
$\sigma(T)$	Spectrum of an operator T
δ	Dirac delta distribution
$p.v.$	Principal value distribution