# MA717: Functional Analysis

Notes

**Anming Gu**
Boston University
agu2002@bu.edu

May 16, 2023

## Abstract

This is a short reference for functional analysis. The material is from Boston University's PhD class in functional analysis, MA717 taught by Mark Kon during Spring 2023. The course follows *Functional Analysis* by Reed and Simon. Topics include theory of Banach and Hilbert spaces, Fourier analysis, wavelet theory and multiresolution analysis, topological vector spaces, ergodic theory, theory of operators, and spectral theory.

## Contents

# 1 Review

## 1.1 Sets and Functions

If $X, Y$ are sets, then $X \times Y = \{(x,y) \mid x \in X, y \in Y\}$ is the *Cartesian product* of $X$ and $Y$.

Let $f : X \to Y$ be a function. The set of all $x$ to which $f$ can be applied is the *domain* of $f$, and the set of all $y$ s.t. $y = f(x)$ for some $x \in X$ is the *range* of $f$. $f$ is *surjective* if $\text{Range}(f) = Y$. $f$ is *injective* if $f(x_1) = f(x_2) \implies x_1 = x_2$. $f$ is *bijective* if it is surjective and injective.

Let $X, Y$ be two sets. A *relation* between $X$ and $Y$ is a subset $R \subset X \times Y$. If $(x,y) \in R$, we say $xRy$ or $x$ is related to $y$. A relation $R \subset X \times X$ is an *equivalence relation* if

(i) $xRx \ \forall x \in X$ (reflexive property);

(ii) $xRy \implies yRx \ \forall x, y \in X$ (symmetric property);

(iii) $xRy, yRz \implies xRz \ \forall x, y, z \in X$ (transitive property).

The set of all elements related to a fixed element $x$ is an *equivalence class.*

**Theorem 1.1.** *Let $R$ be an equivalence relation on a set $X$. Then each element of $X$ belongs to a unique equivalence class.*

## 1.2 Metric Spaces

Let $X$ be a set. Define a *metric* between 2 elements $x, y \in X$ by $\rho(x,y)$ with the following properties

(i) $\rho(x,y) \geq 0 \ \forall x, y \in X$ with equality iff $x = y$ (positivity);

(ii) $\rho(x,y) = \rho(y,x) \ \forall x, y \in X$ (symmetry);

(iii) $\rho(x,y) + \rho(y,z) \geq \rho(x,z) \ \forall x, y, z \in X$ (triangle inequality).

Then $(X, \rho)$ is a *metric space.*

**Example 1.2.** Let $X = C[0,1]$ be the continuous functions on $[0,1]$ and $f_1, f_2 \in X$. Then

$$\rho_1(f_1, f_2) = \int_0^1 |f_1 - f_2| dx,$$
$$\rho_2(f_1, f_2) = \sup_{x\in[0,1]} |f_1(x) - f_2(x)|,$$
$$\rho_3(f_1, f_2) = \left(\int_0^1 |f_1 - f_2|^2 dx\right)^{1/2}$$

are all metrics.

From now on, always assume $(X,\rho), (Y,d)$ are our canonical metric spaces, unless otherwise stated.

Let $\{x_n\}$ be a sequence. Let $x \in X$. Then $x_n \to x$ ($x_n$ *converges* to $x$) if $\rho(x_n, x) \to 0$. If $\forall \varepsilon > 0$, $\exists N$ s.t. $\forall n, m > N$ implies $\rho(x_n, x_m) < \varepsilon$, then $\{x_n\}$ is a *Cauchy* sequence.

**Proposition 1.3.** *If $x_n \in X$ and $\exists x \in X$ s.t. $x_n \to x$, then $\{x_n\}$ is a Cauchy sequence.*

If every Cauchy sequence in $X$ converges to some element of $X$, then $X$ is *complete*.

**Example 1.4.** $\mathbb{R}$ is complete with $\rho(x,y) = |x-y|$, but $\mathbb{Q}$ is not complete with this metric.

$X_1 \subset X$ is *dense* in $X$, if $\forall x \in X, \forall \varepsilon > 0$, $\exists x_1 \in X_1$ s.t. $\rho(x, x_1) < \varepsilon$.

**Example 1.5.** $\mathbb{Q} \subset \mathbb{R}$ is dense in $\mathbb{R}$.

Let $f : X \to Y$ be a bijective map. Assume that $d(f(x), f(y)) = \rho(x,y)$ $\forall x, y \in X$. Then $f$ is called an *isometry*.

**Theorem 1.6.** *If $(X,\rho)$ is not a complete metric space, there exists a complete space $(Y,d)$ s.t. $(X,\rho)$ can be embedded in $(Y,d)$. $Y$ is called the* completion *of $X$.*

**Example 1.7.** $\mathbb{R}$ is the completion of $\mathbb{Q}$.

A function $f : X \to Y$ is *continuous* at $x \in X$ if $\forall \varepsilon > 0$, $\exists \delta > 0$ s.t. if $\rho(x_0, x) < \delta$, then $d(f(x_0), f(x)) < \varepsilon$. $f$ is *continuous* if $f$ is continuous at all $x \in X$.

**Theorem 1.8.** *$f : X \to Y$ is continuous at $x$ iff for any sequence $x_i \to x$ implies $f(x_i) \to f(x)$.*

Let $x \in X$. Then $B_\varepsilon(x) = \{y \mid \rho(x,y) < \varepsilon\}$ is the *open ball of radius* $\varepsilon$ about $y$. A set $\mathcal{O} \subset X$ is *open* if $\forall x \in \mathcal{O}$, $\exists \varepsilon > 0$ s.t. $B_\varepsilon(x) \subset \mathcal{O}$. A *neighborhood* $N$ of $x$ is any set which contains an open set that contains $x$.

Let $E \subset X$. Then $x$ is a *limit point* of $E$ if $\forall r > 0, B_r(x)$ contains points in $E$ besides $x$ itself. $E$ is *closed* if it contains all its limit points.

Given $E \subset X$, we define the *interior* $E^o$ of $E$ to be the largest open set contained in $E$. We define the *closure* $\overline{E}$ of $E$ to be the smallest closed set containing $E$. Then the *boundary* of $E$ is the difference between the two sets: $\partial E = \overline{E} \setminus E^o$. $x \in X$ is a *point of closure* of a set $E$ if $x \in \overline{E}$.

**Proposition 1.9.** *Every point on $\partial\mathcal{O}$ is a limit point of $\mathcal{O}$.*

**Theorem 1.10.** *A set $\mathcal{O} \subset X$ is open iff $\mathcal{O}^c$ is closed.*

**Theorem 1.11.** *$f : X \to Y$ is continuous iff for all open $\mathcal{O} \subset Y$, $f^{-1}(\mathcal{O})$ is open.*[1]

### 1.3 Vector Spaces and Normed Linear Spaces

Recall a *vector space* (v.s.) $V$ is a set that is closed under addition and scalar multiplication that satisfies the following properties:

(a) $\alpha(v_1 + v_2) = \alpha v_1 + \alpha v_2$ (distributive);

(b) $v_1 + v_2 = v_2 + v_1$ (commutative);

(c) $(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$ (associative);

(d) $\alpha_1(\alpha_2 v) = (\alpha_1\alpha_2)v$;

(e) $\exists! v = 0$ s.t. $0 + v_1 = v_1$ $\forall v_1 \in V$;

[1]Note that $f^{-1}$ is the pre-image of $f$, not the inverse function.

(f) $(\alpha_1 + \alpha_2)v = \alpha_1 v + \alpha_2 v$;

(g) $1 \cdot v = v \ \forall v \in V$;

(h) $\forall v, \ \exists w = -v$ s.t. $x + v = 0$.

A v.s. $V$ is a *normed linear space* (n.l.s) if each $v \in V$ is assigned a length $\|v\| \geq 0$ s.t.

(a) $\|v\| = 0$ iff $v = 0$;

(b) $\|\alpha v\| = |\alpha|\|v\| \ \forall \alpha \in \mathbb{R}, v \in V$;

(c) $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|$.

**Remark 1.12.** A n.l.s. is also a metric space if we define $\rho(v_1, v_2) = \|v_1 - v_2\|$.

Let $V_1, V_2$ be n.l.s. Let $T : V_1 \to V_2$. Then $T$ is a *bounded linear transformation* (BLT) if

(i) $T(v_1 + v_2) = Tv_1 + Tv_2$;

(ii) $T(\alpha v_1) = \alpha T v_1$;

(iii) $\|Tv_1\| \leq C\|v_1\| \ \forall v_1 \in V$ for some $C$.

The smallest $C$ which makes (iii) true is called $\|T\|$, the norm of $T$.

**Theorem 1.13.** *Let $T : V_1 \to V_2$ be a linear transformation. Then $T$ is bounded iff $T$ is continuous.*

### 1.4 Countable Sets

A set $A$ is *countable* if there exists a one-to-one mapping between $A$ and $\mathbb{N}$. If $A$ is infinite, and $A$ is not countable, then it is *uncountable*.

**Example 1.14.** $\mathbb{N}, \mathbb{Q}, \mathbb{N}^{100}$ are countable, $\mathbb{R}$ is uncountable.

### 1.5 Measure Theory on $\mathbb{R}$

Motivation: we want to extend the sets that are measurable, *i.e.*, sets that have definite size, from just the intervals on $\mathbb{R}$ to a larger class of subsets of $\mathbb{R}$.

A family $\mathcal{F}$ of subsets of a set $S$ is a $\sigma$-algebra if it contains $S$, it is closed under complements, and it is closed under countable unions. The *Borel sets* $\mathcal{B}$ on $\mathbb{R}$ is the smallest $\sigma$-algebra containing all open intervals in $\mathbb{R}$.

Let $\mathcal{J}$ be the collection of all countable infinite unions of disjoint open intervals in $\mathbb{R}$. For $E \in \mathcal{J}$, write $E = \cup(a_i, b_i)$, where the intervals are assumed to be disjoint. Then the measure of $E$ is $\mu(E) = \sum(b_i - a_i)$. For $B \in \mathcal{B}$, define

$$\mu(B) = \inf_{E \in \mathcal{J} : E \supseteq B} \mu(E).$$

The measure $\mu(B)$ is the *Lebesgue measure* of $B$.

**Theorem 1.15.** *Let $\mu$ be any nonnegative function defined on a $\sigma$-algebra $\mathcal{F}$. $\mu$ is a* measure *if it has the following properties:*

*(i)* $\mu(\varnothing) = 0$

*(ii) If $\{A_n\} \subset \mathcal{F}$, and all $A_n$ are disjoint, then $\mu(\cup A_n) = \sum \mu(A_n)$.*

$f : \mathbb{R} \to \mathbb{R}$ is a *Borel function* if $f^{-1}((a, b)) \in \mathcal{B}$ for any interval $(a, b) \subset \mathbb{R}$.

**Remark 1.16.** Borel functions will be the largest class of functions on which we will be able to define an integral.

**Proposition 1.17.** *$f$ is a Borel function iff $f^{-1}(B) \in \mathcal{B} \ \forall B \in \mathcal{B}$.*

**Theorem 1.18** (Properties of Borel functions)**.** *We have the following:*

*(a) If $f, g$ are Borel measurable, $\lambda \in \mathbb{R}$, then so are $f + g, fg, \max\{f, g\}, \min\{f, g\}, \lambda f$.*

*(b) If $\{f_n\}$ is a sequence of Borel measurable functions, and $f_n \to f$, then $f$ is Borel measurable.*

First, assume that $f$ is nonnegative on $[a, b]$. Fix $n \in \mathbb{N}$. Then for $0 \leq m < \infty$, divide the $y$-axis into intervals $(\frac{m}{n}, \frac{m+1}{n})$. Then we can approximate the integral of $f$, and define the *Lebesgue integral* of $f$ as

$$\int_a^b f = \lim_{n\to\infty} \sum_{m=0}^{\infty} \frac{m}{n} \cdot \mu\left(f^{-1}\left(\frac{m}{n}, \frac{m+1}{n}\right)\right).$$

We will omit the bounds of integration if they are obvious. If $f$ is not positive, then write $f = f_+ - f_-$, where $f_+ = \max\{f, 0\}, f_- = -\min\{f, 0\}$. Now if $\int |f| < \infty$, define

$$\int f = \int f_+ - \int f_-.$$

Define $L^1(a, b) = \{f \mid \int_a^b |f| < \infty\}$.

**Theorem 1.19** (Properties of $L^1$ integrals)**.** *Let $a, b \in \mathbb{R}$ (may be infinite).*

*(a) If $f, g \in L^1(a, b)$, then so are $f + g, \lambda f$ if $\lambda \in \mathbb{R}$. We have the following:*

*(b) If $f \in L^1(a, b)$ and $|g| \leq |f|$, then $g \in L^1(a, b)$.*

*(c) $\int_a^b (f + g) = \int_a^b f + \int_a^b g$.*

*(d) $|\int_a^b f| \leq \int_a^b |f|$.*

*(e) $f, g \in L^1(a, b), f \leq g \implies \int_a^b f \leq \int_a^b g$.*

*(f) $|\int_a^b f| \leq (b - a)\sup_{x\in(a,b)} |f(x)|$.*

**Theorem 1.20** (Monotone convergence theorem)**.** *Let $f_n(x) \geq 0$ be measurable and $f_n(x) \to f(x)$ $\forall x$. Assume $f_{n+1}(x) \geq f_n(x)$ $\forall x, n \in \mathbb{N}$. Then if $\int f_n \leq C$ $\forall n \in \mathbb{N}$, where $C$ is a constant, we have $f \in L^1$,*

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

*and*

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

**Theorem 1.21** (Dominated convergence theorem)**.** *Suppose $|f_n(x)| \leq F(x)$ and $F \in L^1$. Then if $f_n(x) \to f(x)$ $\forall x$, we have $f \in L^1$,*

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

*and*

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

A condition $C$ holds *almost everywhere* (a.e.) on $\mathbb{R}$ if the set of numbers where $C$ is false has measure 0. Functions $f, g$ are *equivalent* ($f \equiv g$) if they differ on a set of measure 0.

**Remark 1.22.** We consider functions which are equivalent to be the same element of $L^1(a, b)$.

Notice that $L^1(a, b)$ is a metric space with metric

$$\rho(f_1, f_2) = \int_a^b |f_1 - f_2| = \|f_1 - f_2\|_1.$$

**Theorem 1.23** (Riesz-Fisher theorem)**.** *$L^1(a, b)$ is complete.*

**Proposition 1.24.** *$C[a, b]$ is a dense subset of $L^1[a, b]$ (in the above metric).*

Recall $\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}$, where $i = \sqrt{-1}$. Let $\alpha = a + bi \in \mathbb{C}$, define the *norm* $|\alpha|$ as $\sqrt{a^2 + b^2}$. We define the *complex conjugate* of $\alpha$ as $\overline{\alpha} = a - bi$.

Now consider $f : \mathbb{R} \to \mathbb{C}$. Then we can write $f = f_1 + if_2$, where $f_1, f_2 : \mathbb{R} \to \mathbb{R}$. Let $|f(x)| = \sqrt{f_1^2(x) + f_2^2(x)}$. Then we have $\|f\|_1 = \int |f|$ and $\int f = \int f_1 + i \int f_2$.

### 1.6 General Measure Theory on $\mathbb{R}$

Let $\alpha : \mathbb{R} \to \mathbb{R}$ be an increasing function (with possible jumps). For an interval $[a, b]$, define the measure $\mu_\alpha([a, b]) = \alpha(b^+) - \alpha(a^-)$. We can show that $\mu_\alpha$ extends uniquely from intervals $[a, b]$ to a measure on $\mathcal{B}$. The measure $\mu_\alpha$ is called the *Stieltjes measure* corresponding to the function $\alpha$. A measure on Borel sets in $\mathbb{R}$ is called a *Borel measure*.

We will use the notation

$$\int f d\mu_\alpha = \int f(x) d\alpha(x).$$

For a function $f : \mathbb{R} \to \mathbb{R}$, we define $f \in L^1(a, b; \mu_\alpha)$ if $\int |f| d\mu_\alpha < \infty$.

In general, if we can write

$$\int f d\mu_\alpha = \int f(x) g(x) dx$$

for some $g(x)$, then $\mu_\alpha$ is called *absolutely continuous* w.r.t. Lebesgue measure.

**Example 1.25.** Let $\alpha(x) = \begin{cases} 1, & x \geq 0; \\ 0, & x < 0. \end{cases}$ Then $\mu_\alpha([a, b]) = \alpha(b^+) - \alpha(a^-) = \begin{cases} 1 & \text{if } 0 \in [a, b]; \\ 0 & \text{if } 0 \notin [a, b], \end{cases}$ *i.e.*, $\{0\}$ has measure 1, and all else has measure 0.

If a measure $\mu$ on $\mathbb{R}$ has a *point mass* at a point $x$ if $\mu(\{x\}) > 0$. If the measure $\mu_\alpha$ is concentrated on a countable or finite set of points, $\mu_\alpha$ is called a *point measure*.

### 1.7 The Cantor Set and the Cantor Function

Consider the set $S$ on $[0, 1]$ such that $S = \left(\frac{1}{3}, \frac{2}{3}\right) \cup \left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right) \cup \left(\frac{1}{27}, \frac{2}{27}\right) \cup \cdots$, *i.e.*, at each stage take the middle third of each interval that's not yet included and include it in the list. Note that $\mu(S) = 1$. Then define $C = [0, 1] \setminus S$. Then $C$ is the *Cantor set*. Note that $\mu(C) = 1 - 1 = 0$.

Now define the *Cantor function* $\alpha(x)$ on $[0, 1]$ as follows. Start with $\alpha(0) = 0$ and $\alpha(1) = 1$, then fill in values on the set $S$ as follows:

$$\alpha(x) = \begin{cases} \frac{1}{2}, & x \in \left(\frac{1}{3}, \frac{2}{3}\right) \\ \frac{1}{4}, & x \in \left(\frac{1}{9}, \frac{2}{9}\right) \\ \frac{3}{4}, & x \in \left(\frac{7}{9}, \frac{8}{9}\right) \\ \text{etc.} \end{cases}$$

We still need to define $\alpha(x)$ on $C$. We can extend $\alpha(x)$ uniquely to a continuous function on $[0, 1]$ by filling in values on $C$ using continuity. That is, for $x \in C$, $\alpha(x) = \lim_{y \uparrow x} \alpha(y)$, where $y \uparrow x$ are chosen so that $y \in S$.

**Lemma 1.26.** *The function $\alpha(x)$ is continuous.*

**Remark 1.27.** $\alpha(x)$ is constant on all intervals forming $S$, meaning $\mu_\alpha(E) = 0$ $\forall E \subset S$. Hence, $\mu_\alpha$ is concentrated on $C$, a set of Lebesgue measure 0! This means that $\mu_\alpha(S) = 0$ while $\mu_\ell(C) = 0$.

We state that $\mu_\alpha$ is *singular* w.r.t. Lebesgue measure $\mu_\ell$ ($\mu_\alpha \perp \mu_\ell$) on $[a, b]$ because there are two sets $S, C$ with $[a, b] = S \cup C$ and $\mu_\alpha(S) = 0$ and $\mu_\ell(C) = 0$, *i.e.*, two measure have support in entirely different places.

A function $\alpha : \mathbb{R} \to \mathbb{R}$ is *right continuous* if its value at any point is the limit of its values to the right, *i.e.*, $\alpha(x) = \lim_{y \downarrow x} \alpha(y)$, where the limit $y \downarrow x$ indicates that $y$ approaches $x$ from the right.

**Proposition 1.28.** *There is a one-to-one correspondence between right continuous functions $\alpha(x)$ on the interval $(a, b)$ and Borel measure $\mu(x)$.*

A Stieltjes measure $\mu_\alpha$ arising from a continuous function $\alpha(x)$ is called a *continuous measure*. A continuous measure $\mu_\alpha$ which is singular w.r.t. Lebesgue measure is called a *singular continuous measure*.

**Proposition 1.29.** *A measure $\mu$ is continuous iff it has no point masses.*

### 1.8 General Measures on Borel Sets

Let $\mu$ be a measure on $\mathcal{B}$ in $\mathbb{R}$. Assume also that $\forall B \in \mathcal{B}$,

$$\begin{aligned} \mu(B) &= \inf_{\substack{\mathcal{O} \supset B \\ \mathcal{O} \text{ open}}} \{\mu(\mathcal{O})\} && (1) \\ &= \sup_{\substack{\mathcal{C} \subset B \\ \mathcal{C} \text{ closed, bdd}}} \{\mu(\mathcal{C})\}, && (2) \end{aligned}$$

then $\mu$ is called a *Borel measure*. If Property 1 holds for all Borel sets, then $\mu$ is called *outer regular*; if Property 2 holds for all Borel sets, then $\mu$ is called *inner regular*.

**Theorem 1.30.** *Any Borel measure $\mu$ can be written as a sum $\mu = \mu_p + \mu_c$, where $\mu_c$ is a continuous measure and $\mu_p$ is a pure point measure.*

There is another way to break down all Borel measure on $\mathbb{R}$.

**Theorem 1.31** (Lebesgue decomposition theorem)**.** *Any Borel measure $\mu$ on $\mathbb{R}$ can be uniquely written as a sum $\mu = \mu_{ac} + \mu_s$, where $\mu_{ac}$ is absolutely continuous w.r.t. Lebesgue measure and $\mu_s$ is singular w.r.t. Lebesgue measure.*

**Theorem 1.32.** *Any Borel measure $\mu$ on $\mathbb{R}$ can be written as $\mu = \mu_p + \mu_{ac} + \mu_{cs}$, where $\mu_p$ is a pure point, $\mu_{ac}$ is absolutely continuous w.r.t. Lebesgue measure, and $\mu_{cs}$ is singular w.r.t. Lebesgue measure.*

Note that *cs* stands for *continuous singular*.

### 1.9 Abstract Measure Theory

Let $(M, \mathcal{F}, \mu)$ be our canonical measure space unless otherwise stated.

A set $A \in \mathcal{F}$ is called *measurable*. A measure $\mu$ is a *$\sigma$-finite* measure on $M$ if $M$ can be written as a countable union of sets $A_i$ s.t. $\mu(A_i) < \infty\ \forall i$.

**Remark 1.33.** We will often assume that our measures are $\sigma$-finite; otherwise things can get too large.

Let $f : M \to \mathbb{R}$. Then we say $f$ is *measurable* if $f^{-1}((a,b))$ is a measurable set for any interval $(a,b)$.

For a measure space $(M, \mathcal{A}, \mu)$ and function $f : M \to \mathbb{R}$ (where $f$ is nonnegative), we define the integral identical in form to the Lebesgue integral on $\mathbb{R}$:

$$\int f\, d\mu = \lim_{n\to\infty} \sum_m \frac{m}{n} \cdot \mu\left(f^{-1}\left(\frac{m}{n}, \frac{m+1}{n}\right)\right).$$

Then to extend this for general real functions, again define $f = f_+ - f_-$. We can also extend integration to complex functions $f : M \to \mathbb{C}$ by writing $f = f_1 + i f_2$ for $f_1, f_2 : M \to \mathbb{R}$. We get exact analogues of the monotone convergence theorem, the dominated convergence theorem, and the Riesz-Fisher theorem.

Let $(M, \mathcal{A}_M, \mu), (N, \mathcal{A}_N, \nu)$ be measure spaces. Let $M \times N$ be the product space. Define the *product $\sigma$-algebra* $\mathcal{A}_{M\times N}$ to be the smallest $\sigma$-algebra on $M \times N$ that contains all product sets of the form $A \times B$ s.t. $A \in \mathcal{A}_M, B \in \mathcal{A}_N$. Sets of this form are called *measurable rectangles*.

**Theorem 1.34.** *If the above two measure $\mu, \nu$ are $\sigma$-finite, there exists a unique measure $\mu \times \nu$ on $\mathcal{A}_{M\times N}$ with the property $(\mu \times \nu)(A \times B) = \mu(A)\nu(B)\ \forall A \in \mathcal{A}_M, B \in \mathcal{A}_N$. This measure is the* product measure *of $\mu, \nu$.*

**Theorem 1.35** (Fubini's theorem)**.** *Assume $(M, \mathcal{A}_M, \mu), (N, \mathcal{A}_N, \nu)$ are $\sigma$-finite. If $f(x,y)$ is a measurable function on $M \times N$, we have*

$$\int_M \left(\int_N f(x,y)\, d\nu(y)\right) d\mu(x) = \int_{M\times N} f(x,y)\, d(\mu\times\nu) = \int_N \left(\int_M f(x,y)\, d\mu(x)\right) d\nu(y),$$

*if any of the three integrals converges absolutely.*

Two measures $\mu, \nu$ on a measurable space $(M, \mathcal{F})$ are *mutually singular* if $\exists A \in \mathcal{A}$ s.t. $\mu(A) = 0$ and $\nu(A^c) = 0$. $\nu$ is *absolutely continuous* w.r.t. $\mu$ ($\nu \ll \mu$) if $\forall A \in \mathcal{F}$, $\nu(A) = 0$ whenever $\mu(A) = 0$.

**Theorem 1.36** (Radon-Nikodym theorem)**.** *$\nu \ll \mu$ on $(M, \mathcal{F})$ iff $\nu(A) = \int_A f\, d\mu$ for some measurable $f$ and $\forall A \in \mathcal{F}$. We call the function $f = \frac{d\nu}{d\mu}$ the* Radon-Nikodym derivative *of $\nu$ w.r.t. $\mu$.*

**Theorem 1.37** (Lebesgue decomposition theorem)**.** *Given two fixed measure $\mu, \nu$, there exists a unique decomposition $\nu = \nu_{ac} + \nu_s$ s.t. $\nu_{ac} \ll \mu$ and $\nu_s \perp \mu$.*

### 1.10 Inner Product Spaces

Let $V$ be a complex v.s. Assume that $\forall v, w \in V$, there exists a number $\langle v, w\rangle \in \mathbb{C}$, called an *inner product* s.t.

(a) $\langle v, v\rangle \geq 0$ with equality iff $x = 0$;

(b) $\langle v, w+u\rangle = \langle v, w\rangle + \langle v, u\rangle$;

(c) $\langle u, \alpha v\rangle = \alpha\langle u, v\rangle$;

(d) $\langle u, v\rangle = \overline{\langle v, u\rangle}$.

Then $V$ is an *inner product space* (i.p.s.).

**Remark 1.38.** We also have $\langle \alpha u, v\rangle = \overline{\alpha}\langle u, v\rangle$.

**Example 1.39.** Let $V = C[0,1]$. Then if $f_1, f_2 \in V$, define $\langle f_1, f_2\rangle = \int \overline{f_1} f_2$.

If $V$ is an i.p.s., $x, y \in V$ are *orthogonal* if $\langle x, y\rangle = 0$. If $\{v_i\}_{i=1}^k \subset V$ ($k$ can be $\infty$) is a set of orthogonal vectors s.t. $\langle v_i, v_i\rangle = 1\ \forall i$, then $\{v_i\}$ is called an *orthonormal set* of vectors. For a vector $v \in V$, define $\|v\| = \sqrt{\langle v, v\rangle}$.

**Theorem 1.40** (Pythagorean theorem)**.** *Let $\{x_i\}_{i=1}^k$ be an orthonormal set ($k < \infty$). Then $\forall x \in V$,*

$$\|x\|^2 = \sum_i \|\langle x, x_i\rangle|^2 + \left\| x - \sum_i \langle x_i, x\rangle x_i \right\|^2 .$$

**Corollary 1.41** (Bessel's inequality)**.** *Let $\{x_i\}_{i=1}^k$ be an orthonormal set ($k < \infty$). Then $\forall x \in V$,*

$$\|x\|^2 \geq \sum_i |\langle x, x_i\rangle|^2.$$

**Corollary 1.42** (Schwarz's inequality)**.** $|\langle x, y\rangle| \leq \|x\|\|y\|\ \forall x, y \in V$.

**Theorem 1.43** (Metric induced by a norm)**.** *If $V$ is an i.p.s. and $\|x\| = \sqrt{\langle x, x\rangle}$, then $V$ with norm $\|x\|$ satisfies the properties of being a n.l.s. Thus, we also have metric $\rho(x, y) = \|x - y\|$.*

## 2 Hilbert Spaces

If i.p.s. $V$ is complete, then $V$ is called a *Hilbert space*. Our canonical Hilbert space will be $\mathcal{H}$.

**Example 2.1.** $L^2[0,1] =$ functions $f : [0,1] \to \mathbb{C}$ s.t. $\int |f|^2 < \infty$ with inner product $\langle f_1, f_2\rangle = \int \overline{f_1} f_2$ is a Hilbert space.

**Example 2.2.** $\ell^2 =$ sequences $\{x_i\}$ of complex numbers s.t. $\sum |x_i|^2 < \infty$ with inner product $\langle \{x_i\}, \{y_i\}\rangle = \sum \overline{x_i} y_i$ is a Hilbert space.

Let $V$ be a Hilbert space. and $M$ be a closed subspace of $V$. Define $M^\perp = \{x \in V \mid x \perp v\ \forall v \in M\}$. Then $M^\perp$ is a closed subspace of $V$.

**Theorem 2.3** (Projection theorem)**.** *Every vector $v \in \mathcal{H}$ can be uniquely written as $v = x + y$, where $x \in M$ and $y \in M^\perp$.*

Let $T : \mathcal{H} \to \mathbb{C}$ be linear and continuous. Then $T$ is a *continuous linear functional.* The collection $\mathcal{H}^*$ of all such $T$'s is the *dual space* of $\mathcal{H}$.

**Theorem 2.4** (Riesz representation theorem)**.** *Let $T : \mathcal{H} \to \mathbb{C}$ be a continuous linear functional. Then $\exists! y \in \mathcal{H}$ s.t. $Tx = \langle y, x\rangle\ \forall x \in \mathcal{H}$. Conversely, $\forall y \in \mathcal{H}$, if we define $Tx = \langle y, x\rangle$, then $T$ is a continuous linear functional.*

Recall that a collection $S = \{x_i\} \subset \mathcal{H}$ is orthonormal if the $x_i$ are orthogonal to each other and have length 1. If $S$ cannot be extended to a larger orthonormal set, then $S$ is an *orthonormal basis* for $\mathcal{H}$.

Let $\{x_\alpha\}_{\alpha \in A}$ be a set. Note that if $A$ is uncountable, a sum $\sum c_\alpha x_\alpha$ is defined (finite) iff at most a countable number of terms in the sum are non-zero. We say that a countable sum $\sum c_i x_i = y$ iff $\lim_{n\to\infty} \| \sum_{i=1}^n c_i x_i - y\| = 0$.

**Remark 2.5.** Now we can continue to talk about Hilbert spaces with uncountable bases.

**Theorem 2.6.** *Let $\{x_\alpha\}_{\alpha \in A}$ be an orthonormal basis (that could be uncountable). Then $\forall y \in \mathcal{H}$, $y = \sum c_\alpha x_\alpha$, where $c_\alpha = \langle x_\alpha, y\rangle$. Further, $\|y\|^2 = \sum |c_\alpha|^2$. Also, if $\sum |c_\alpha|^2 < \infty$ for some collection $\{c_\alpha\}_{\alpha \in A}$ of constants, then $\sum c_\alpha x_\alpha$ converges to an element of $\mathcal{H}$.*

**Remark 2.7.** Note that Hilbert spaces generalize our intuition of finite-dimensional vector spaces from linear algebra.

Let $\{u_i\}$ be linearly independent vectors. We can construct orthonormal vectors $\{v_i\}$ which are special combinations of $\{u_i\}$. Define $w_i, v_i$ by

$$\begin{aligned} w_1 &= u_1 & v_1 &= \tfrac{w_1}{\|w_1\|} \\ w_2 &= u_2 - \langle v_1, u_2\rangle v_1 & v_2 &= \tfrac{w_2}{\|w_2\|} \\ w_3 &= u_3 - \langle v_1, u_3\rangle v_1 - \langle v_2, u_3\rangle v_2 & v_3 &= \tfrac{w_3}{\|w_3\|} \\ &\vdots \end{aligned}$$

This is the *Gram-Schmidt procedure*, which is entirely analogous to the standard procedure from linear algebra.

Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces. Let $U : \mathcal{H}_1 \to \mathcal{H}_2$ be a linear map which is

(i) inner product preserving, *i.e.*, $\langle Ux, Uy\rangle = \langle x, y\rangle$;

(ii) onto, *i.e.*, the range of $U$ is $\mathcal{H}_2$;

(iii) one-to-one, *i.e.*, $v_1 \neq v_2 \implies Uv_1 \neq Uv_2$.

Then $U$ is called a *unitary map.* If there exists a unitary map $U$ from $\mathcal{H}_1$ to $\mathcal{H}_2$, then $\mathcal{H}_1$ and $\mathcal{H}_2$ are *isomorphic* ($\mathcal{H}_1 \sim \mathcal{H}_2$). Alternatively, they can also be called *unitarily equivalent.*

Let $(X, \rho)$ be a metric space. $X$ is *separable* if it has a dense subset $S$ which is countable.

**Theorem 2.8.** *A Hilbert space $\mathcal{H}$ is separable iff there exists an orthonormal basis $\{x_\alpha\}_{\alpha\in A}$ which is countable.*

**Remark 2.9.** We will mainly care about separable Hilbert spaces.

**Theorem 2.10.** *Let $\mathcal{H}$ be a Hilbert space. We have the following:*

*(a) If $\mathcal{H}$ has a finite orthonormal basis $\{x_i\}_{i=1}^n$, then $\mathcal{H} \sim \mathbb{C}^n$.*

*(b) If $\mathcal{H}$ has a countable orthonormal basis $\{x_i\}_{i=1}^\infty$, then $\mathcal{H} \sim \ell^2$.*

### 2.1 Fourier Series

Let $f$ be integrable on $[-\pi, \pi]$. We can show that if $\phi_n(x) = \frac{e^{inx}}{\sqrt{2\pi}}$, then $\{\phi_n(x)\}_{n=-\infty}^\infty$ is orthonormal. Therefore, if $f \in L^2[-\pi, \pi]$, then

$$f(x) = \frac{1}{\sqrt{2\pi}} \sum d_n e^{inx},$$

where $d_n = \langle \phi_n, f\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{-inx} f(x)\, dx$. Alternatively, we can write

$$f(x) = \sum c_n e^{inx},$$

where $c_n = \frac{d_n}{\sqrt{2\pi}}$.

**Remark 2.11.** We call $f_M(x) = \sum_{n=-M}^M c_n e^{inx}$. Then $\lim_{M\to\infty} f_M(x) = f(x)$ in $L^2$.

**Theorem 2.12.** *If $f \in L^2[-\pi, \pi]$, then $f_M \to_{M\to\infty} f$ in $L^2$ norm, i.e., $\|f_M - f\|_2^2 = \int |f_M - f|^2 \xrightarrow{M\to\infty} 0$.*

**Remark 2.13.** $L^2[-\pi, \pi]$ is a Hilbert space.

We can also use Euler's formula to show that the functions consisting of

$$\frac{1}{\sqrt{2\pi}}, \left\{\frac{1}{\sqrt{\pi}} \cos nx\right\}_{n=1}^\infty, \left\{\frac{1}{\sqrt{\pi}} \sin nx\right\}_{n=1}^\infty$$

equivalently forms an orthonormal basis. Similar analysis then gives

$$f(x) = \frac{1}{2} a_0 + \sum_{n=1}^\infty a_n \cos nx + b_n \sin nx,$$

where $a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx\, dx, b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx\, dx$.

**Remark 2.14.** In 1966, Carleson showed that convergence of Fourier series occurs pointwise a.e.

### 2.2 Applications in Quantum Mechanics

There are two classical views of quantum mechanics: Heisenberg's and Schrodinger's.

Heisenberg viewed all possible states of a system forming a vector space, where all states are infinite vectors in a Hilbert space $V_H$. Suppose $v = (x_1, x_2, \dots) \in V_H$. Then $|x_i|^2$ represents the probability that particles are in a particular configuration, with each $i$ representing a different configuration. For example, if $v_1, v_2$ are two possible states, then $\frac{1}{2}v_1 + \frac{1}{2}v_2$ is another possible state. Then the time evolution of $v = v(t)$ is given by a matrix equation

$$\frac{dv}{dt} = Hv,$$

where $H$ is an infinite matrix.

Now let $V_S$ be the set of all functions on some set. For the sake of simplicity, assume $V_S = \{\psi \mid \psi : [-\pi, \pi] \text{ s.t. } \int |\phi|^2 < \infty\}$. We interpret $|\psi(x)|^2 \Delta x$ as the probability a particle is in the interval $\Delta x$. Actually, $\psi = \psi(x,t)$ satisfies a partial differential equation for time evolution:

$$-i\frac{\partial \psi(x,t)}{\partial t} = \frac{\partial^2 \psi(x,t)}{\partial x^2}.$$

Equivalently, we can define for a fixed $t$, $\Psi(t) = \psi(x,t) \in L^2(\mathbb{R})$.

Schrodinger showed that the two views of quantum mechanics were unitarily equivalent, i.e. there exists a bijective unitary operator $U : V_S \to V_H$ such that $U\Psi = v$.

## 3 Banach Spaces

Let $(X, \mu)$ be a measure space. Given a function $f : X \to \mathbb{R}$, we define the *essential supremum* of $f$ ($\operatorname{ess\,sup} f$) to be the maximum value of $|f(x)|$, except for sets of measure 0. Precisely,

$$\operatorname{ess\,sup} f = \inf\{a \mid |f(x)| \leq a \ \forall x \text{ a.e.}\}$$

If a n.l.s. $V$ is complete, then $V$ is called a *Banach space.*

**Example 3.1.** Let $1 \leq p < \infty$, and $(X, \mu)$ be a measure space. Let $L^p = \{f : X \to \mathbb{C} \mid \int |f|^p d\mu < \infty\}$. This is a n.l.s.

**Example 3.2.** If $\mathcal{H}$ is a Hilbert space, then $\mathcal{H}$ can be considered a Banach space using the norm $\|x\| = \sqrt{\langle x, x\rangle}$.

**Theorem 3.3** (Riesz-Fisher theorem). *$L^p(X)$ is complete.*

**Theorem 3.4** (Holder's inequality). *Let $1 \leq p \leq \infty$, and $\frac{1}{p} + \frac{1}{q} = 1$. Let $f \in L^p(X), g \in L^q(X)$. Then $fg \in L^1(X)$, and $\|fg\|_1 \leq \|f\|_p \|g\|_q$.*

**Corollary 3.5.** *If $f \in L^p(X)$ for $1 \leq p \leq \infty$, then if $\frac{1}{p} + \frac{1}{q} = 1$, $\|f\|_p = \sup\{\|fg\|_1 \mid \|g\|_q = 1\}$.*

**Corollary 3.6** (Triangle inequality). *If $1 \leq p \leq \infty$, we have $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.*

Define $L^p(X)^*$ to be the continuous linear functionals on $L^p(X)$.

**Theorem 3.7** (Riesz representation theorem). *Let $(X, \mu)$ be a $\sigma$-finite measure space. Let $1 \leq p < \infty$. Let $L \in L^p(X)^*$. Then $\exists! g \in L^q(X)$ s.t.*

$$L(f) = \int f(x)g(x)\, d\mu(x),$$

*and furthermore, $\|g\|_q = \|L\|$.*

### 3.1 Dual Spaces

Let $X$ be a n.l.s. Let $\{x_n\} \subset X$ be a sequence. Then $\{x_n\}$ is *summable* if $\lim_{N\to\infty} \sum_{n=1}^{N} x_n = \sum x_n$ exists. $\{x_n\}$ is *absolutely summable* if $\sum \|x_n\|$ converges.

**Theorem 3.8.** *$X$ is complete iff every absolutely summable sequence is summable.*

Given a Banach space $X$, the dual space $X^*$ is also a Banach space. To see this, note that $X^* = \mathcal{L}(X, \mathbb{C})$, which is a Banach space.

**Example 3.9.** Let $(M, \mu)$ be a measure space. Then $L^p(M)$ be the functions $f$ on $M$ s.t. $\|f\|_p = (\int |f|^p)^{1/p} < \infty$ is a Banach space under the norm $\|f\|_p$ for $1 \le p < \infty$. Then $\exists g \in L^q(M)$ where $1/q + 1/p = 1$ s.t. $Lf = \int_X \overline{g} f d\mu$, where $\|L\| = \|g\|_q$. Thus, $L \in L^p(M)^* \longleftrightarrow g \in L^q(M)$. Let this correspondence $A$. It is easy to check that $A$ is linear and bijective. Also, the operator norm $\|L\| = \|g\|_q$, so $L$ is isometric. Hence, $L^p(X)^*$ and $L^q(X)$ are the same Banach space under this correspondence.

Two n.l.s. $X, Y$ are *isomorphic* if there exists a linear map $A : X \to Y$ that is bijective and isometric. This means that $X, Y$ are effectively the same space.

**Example 3.10.** Let $\mathcal{H}$ be a Hilbert space. We have shown that if $L \in \mathcal{H}^*$, then $\exists y \in \mathcal{H}$ s.t. $Lx = \langle y, x \rangle \ \forall x$. Thus, we have an analogous correspondence $M$ where $L \in \mathcal{H}^* \longleftrightarrow y \in \mathcal{H}$. $M$ forms a bijection between $\mathcal{H}^*, \mathcal{H}$, $\|L\|_{\mathcal{H}^*} = \|y\|_{\mathcal{H}}$, so $M$ is an isomorphism. Hence, $\mathcal{H}^*, \mathcal{H}$ are isomorphic.

**Claim 3.11.** *Let* $\ell_1 = \{\{a_i\} \mid \sum |a_i|^1 < \infty\}$ *and* $\ell_\infty = \{\lambda = \{\lambda_i\} \mid \|\lambda\|_\infty = \sup |\lambda_i| < \infty\}$. *Then* $\ell_1^* = \ell_\infty$.

Let $X$ be a Banach space. Then $X^{**}$ is the space of linear functionals on $X^*$, the double dual space.

**Theorem 3.12.** *Let $X$ be a Banach space and $X^*$ be its dual space. Fix $x \in X$, and define the following linear functional $\tilde{x}$ defined on $\lambda \in X^*$ by $\tilde{x}(\lambda) = \lambda(x)$. Define the map $J : X \to X^{**}$ by $Jx = \tilde{x}$. Then $J$ is an isomorphism of $X$ onto some subspace of $X^{**}$ (possibly all of $X^{**}$).*

If $J$ maps $X$ onto all of $X^{**}$, we say $X$ is reflexive.

### 3.2 The Hahn-Banach Theorem

Let $S$ be a set. A relation $\prec$ on $S$ is a *partial order* if:

(a) $x \prec x \ \forall x \in S$

(b) $x \prec y, y \prec z \implies x \prec z$

(c) $x \prec y, y \prec x \implies x = y$.

If $\forall x, y \in S$, x$\prec$ $y$ or $y \prec x$, $s$ is *linearly ordered.* Suppose $X \subset S$, and $p \in S$ satisfies the condition that $x \prec y \ \forall x \in X$. Then we say $p$ is an *upper bound* for $X$. Suppose that $m \in X$, and $\nexists x \neq m$ s.t. $m \prec x$, then we say that $m$ is a *maximal element* in $S$.

**Lemma 3.13** (Zorn's lemma)**.** *Let $S$ be a set with a partial order, s.t. every linearly ordered subset $X \subset S$ has an upper bound $p$. Then $X$ has a maximal element in $S$.*

**Theorem 3.14** (Hahn-Banach theorem)**.** *Let $X$ be a vector space over the real numbers. Let $p : X \to \mathbb{R}$ be a function s.t. $\forall \alpha \in [0, 1], x \in X$, $p(\alpha x + (1 - \alpha)y) \le \alpha p(x) + (1 - \alpha)p(y)$. Let $\lambda$ be a linear functional defined on subspace $Y \subset X$ s.t. $\lambda(x) \le p(x) \ \forall x \in Y$. Then there is a linear functional $\Lambda : X \to \mathbb{R}$ s.t. $\Lambda(x) = \lambda(x) \ \forall x \in Y$ and $\Lambda(x) \le p(x) \ \forall x \in X$.*

**Corollary 3.15.** *Let $X$ be a n.l.s. Let $Y \subset X$ be a subspace, and $\lambda$ be a bounded linear functional on $Y$. Then there exists a bounded linear functional $\Lambda$ on $X$ which extends $\lambda$ s.t. $\|\Lambda\| = \|\lambda\|$.*

### 3.3 Main Theorems on Banach Spaces

Recall if $M$ is a metric space, $A \subset M$, then $A^o$ is the interior of $A$, *i.e.*, the largest open set contained in $A$. A set $A \subset M$ is *nowhere dense* if $\overline{A}$ has an empty interior, *i.e.*, $\overline{A}$ contains no open set.

**Example 3.16.** Suppose $A = \{x_i\}_{i=1}^n \subset \mathbb{R}$, *i.e.*, a finite sequence. Then $A$ is nowhere dense since $\overline{A} = A$, and $A$ contains no open balls.

**Example 3.17.** $\mathbb{Q} \subset R$ is not nowhere dense because $\overline{\mathbb{Q}} = \mathbb{R}$ and $\mathbb{R}$ does contain open balls.

**Proposition 3.18.** *Let $X, Y$ be n.l.s.'s. Let $T : X \to Y$ be linear. Then $T$ is bounded iff $T^{-1}(B)$ has nonempty interior, where $B = \{y \in Y \mid \|y\| \le 1\}$ is the closed unit ball of $Y$.*

Recall that in a metric space $X$, $A \subset X$ is *first category* if $A = \cup A_n$ s.t. $A_n$ are nowhere dense. $A$ is *second category* if it is not first category.

**Theorem 3.19** (Baire category theorem)**.** *A complete metric space is never the union of a countable number of nowhere dense sets, i.e., it is second category.*

**Theorem 3.20** (Banach-Steinhaus theorem)**.** *Let $X$ be a Banach space, $Y$ be a n.l.s., and $\mathcal{F}$ be a family of bounded linear transformations from $X$ to $Y$. Suppose that $\forall x \in X$, $S_x = \{\|Tx\| \mid T \in \mathcal{F}\}$ is a bounded set* *in* $\mathbb{R}$. *Then the set*

$$S = \{\|T\| \mid T \in \mathcal{F}\}$$

*is also bounded.*

**Theorem 3.21** (Open mapping theorem)**.** *Let* $T : X \to Y$ *be a bounded linear transformation, where* $X, Y$ *are Banach spaces. If* $E$ *is an open set in* $X$*, then* $T(E)$ *is an open set in* $Y$*.*

**Corollary 3.22** (Inverse mapping theorem)**.** *Let* $X, Y$ *be Banach spaces. Let* $T : X \to Y$ *be a linear transformation which is continuous and a bijection. Then* $T^{-1}$ *is continuous.*

Let $X, Y$ be n.l.s's. Define $Z = X \oplus Y = \{(x, y) \mid x \in X, y \in Y\}$ with the following properties:

(i) If $z_1 = (x_1, y_1), z_2 = (x_2, y_2)$, then $z_1 + z_2 = (x_1 + x_2, y_1 + y_2)$.

(ii) If $c \in \mathbb{C}, z = (x, y)$, then $cz = (cx, cy)$.

We can show $Z$ is a vector space. If $z = (x, y)$, then define the norm $\|z\| = \|x\| + \|y\|$, so $Z$ is a n.l.s. If $X, Y$ are Banach spaces, then $Z$ also is a Banach space. Let $T : X \to Y$. Then the *graph* of $T$ is $G = \{(x, Tx) \mid x \in X\} \subset Z$.

**Theorem 3.23** (Closed graph theorem)**.** *Let* $X, Y$ *be Banach spaces,* $Z = X \oplus Y$*, and* $T : X \to Y$ *be a linear map. Then* $T$ *is bounded iff the graph of* $T$ *is a closed set in* $Z$*.*

# 4 Fourier Analysis and Wavelet Theory

## 4.1 Basic Properties of the Fourier Transform

Let $\mathcal{F}$ denote the Fourier transform operator. For a function $f \in L^2(-\infty, \infty)$, we can write

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} c(\xi) e^{ix\omega}\, d\omega,$$

where $\hat{f}(\omega)$, called the *Fourier transform* of $f$ is

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ix\omega}\, dx.$$

Consider a discontinuous function $f$ and try to apply the first few partial sums of its Fourier series. It turns out that a single discontinuity results in large errors appearing near the singularity, which propagates the error elsewhere too! In general, singularities cause high-frequency components so that the Fourier series has large $c_n$ for large $n$, which is bad for convergence. Wavelets can deal with the problem of localization of singularities.

There are several advantages of Fourier series. Frequency content are displayed in sizes of coefficients $a_n$ and $b_n$. It is also easy to write derivatives of $f$ in terms of series. However, Fourier series is usually not well-adapted for time-frequency analysis.

**Theorem 4.1** (Plancherel's theorem)**.** *We have the following:*

*(i) The Fourier transform is a one-to-one correspondence from* $L^2$ *to itself.*

*(ii) The Fourier transform preserves inner products, i.e., if* $\hat{f}, \hat{g}$ *are the Fourier transforms of* $f, g$*, respectively, then* $\langle \hat{f}(\omega), \hat{g}(\omega)\rangle = \langle f(x), g(x)\rangle$*.*

*(iii) Thus,* $\|f(x)\|^2 = \|\hat{f}(\omega)\|^2$*.*

If we the Fourier series of a function $f \in L^2[-\pi, \pi]$, the above theorem has an analog on $[-\pi, \pi]$:

**Theorem 4.2** (Plancherel's theorem for Fourier series)**.** *We have the following:*

*(i) The correspondence between functions* $f \in L^2[-\pi, \pi]$ *and the coefficients* $\{c_k\}$ *of their Fourier series is a one-to-one correspondence, if we restrict* $\sum c_k^2 < \infty$*.*

*(ii) Furthermore,* $\sum \|c_k\|^2 = \frac{1}{2\pi}\|f(x)\|^2$*.*

Th *convolution* of two functions $f(x)$ and $g(x)$ is defined to be

$$(f * g)(x) = \int_{-\infty}^{\infty} f(x - y) g(y)\, dy.$$

**Proposition 4.3** (Properties of the Fourier transform)**.** *Let $\hat{f} = \mathcal{F}(f)$. Then, the Fourier transform has the following properties:*

*(i)* $\mathcal{F}(f(x-c)) = e^{-i\omega c}\hat{f}(\omega)$.

*(ii)* $\mathcal{F}(f'(x)) = i\omega\hat{f}(\omega)$.

*(iii)* $\mathcal{F}(f(cx)) = \frac{1}{c}\hat{f}(\omega/c)$.

*(iv)* $\mathcal{F}(xf(x)) = i[\hat{f}(\omega)]'$.

*(v)* $\mathcal{F}((f * g)(x)) = \sqrt{2\pi}\hat{f}(\omega)\hat{g}(\omega)$.

*(vi)* $\mathcal{F}(f(-x)) = \overline{\hat{f}}(\omega)$.

**Lemma 4.4.** *The Fourier transform of an integrable function is absolutely continuous.*

**Theorem 4.5** (Decay rate of the Fourier transform)**.** *We have the following:*

*(i) If a function $\psi(x)$ has $n$ derivatives which are integrable and which go to 0 at $\infty$, then the Fourier transform satisfies*

$$|\hat{\psi}(\omega)| \leq K(1+|\omega|)^{-n}. \tag{3}$$

*Conversely, if Equation (3) holds, then $\psi(x)$ has at least $n-2$ derivatives.*

*(ii) Similarly, if $\hat{\psi}(\omega)$ and its first $n$ derivatives are integrable and go to 0 at $\infty$, then*

$$|\psi(x)| \leq K(1+|x|)^{-n}. \tag{4}$$

*Conversely, if Equation (4) holds, then $\hat{\psi}(\omega)$ has at least $n-2$ derivatives.*

### 4.2 Wavelet Transform

Consider a fixed function $h(x)$. Then form all translations by integers, and all scalings by powers of 2:

$$h_{jk}(x) = 2^{j/2}h(2^j x - k).$$

Then let

$$c_{jk} = \int f(x)h_{jk}(x)\,dx = \langle f, h_{jk}\rangle.$$

If $h$ is chosen properly, then we can get back $f$ from the $c_{jk}$:

$$f(x) = \sum_{j,k} c_{jk}h_{jk}(x).$$

We will show that it is possible to find a function $h$ s.t. the functions $h_{jk}$ form a perfect basis for functions on $\mathbb{R}$, that is, the $h_{jk}$ are orthogonal and any function can be represented by the $h_{jk}$. As a result, wavelet series are like Fourier series, but $h_{jk}$ are better, *e.g.*, non-zero only on a small sub-interval, compactly supported.

### 4.3 Haar Wavelets

Consider basis function $\phi(x) = \chi_{[0,1]}$, the basic pixel. We want to build all other functions out of this and translates $\phi(x-k)$ for $k \in \mathbb{Z}$. Define $V_0$ to be all square integrable functions of the form $g(x) = \sum_k a_k\phi(x-k)$, *i.e.*, the square integrable functions which are constant on integer intervals. To get better approximations, we shrink the pixels. Define $V_j$ to be all square integrable functions of the form $g(x) = \sum_k a_k\phi(2^j x - k)$, *i.e.*, the square-integrable functions which are constant on $2^{-j}$ length intervals.

Then define the wavelet

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

We define the family of Haar wavelets by translating and stretching the basic wavelet above. In general, we have

$$\psi_{jk} := 2^{j/2}\psi(2^j x - k).$$

It is easy to show that the $\psi_{jk}$ are orthogonal. Can any function be represented as a combination of Haar wavelets? Recall the definition of $V_j$. Notice that we have

(a) $\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots$.

(b) $\cap V_n = \{0\}$.

(c) $\cup V_n$ is dense in $L^2(\mathbb{R})$.

(d) $f(x) \in V_n \implies f(2x) \in V_{n+1}$.

(e) $f(x) \in V_0 \implies f(x-k) \in V_0$.

(f) There exists an orthogonal basis for the space $V_0$ in the family of functions $\phi_{0k} := \phi(x-k)$ where $k \in \mathbb{Z}$. This function $\phi$ is (in this case) $\phi = \chi_{[0,1]}(x)$. $\phi$ is called a *scaling function*.

A sequence of spaces $\{V_j\}$ together with a scaling function $\phi$ which generates $V_0$ s.t. (a)-(f) above are satisfied is called a *multiresolution analysis* (MRA).

Recall that a vector space $V$ is a *direct sum* $M_1 \oplus M_2$ of subspaces $M_1, M_2$ if every vector $v \in V$ can be written uniquely as a sum of vectors $w_1 \in M_1$ and $w_2 \in M_2$. $V$ is an *orthogonal* direct sum $M_1 \oplus M_2$ if the above holds and in addition $M_1$ and $M_2$ are orthogonal.

**Theorem 4.6.** *Assume $V$ is a Hilbert space and subspaces $M_1 \perp M_2$. Assume also that $V = M_1 + M_2$, i.e., $\forall v \in V$, $\exists m_i \in M_i$ s.t. $v = m_1 + m_2$. Then $V = M_1 \oplus M_2$ is an orthogonal direct sum of $M_1$ and $M_2$.*

If $V = W_1 \oplus W_2$ is an orthogonal direct sum, we also write $W_1 = V \ominus W_2$ and $W_2 = V \ominus W_1$.

Recall the definition of $V_j$. Since $V_0 \subset V_1$, there exists a subspace $W_0 = V_1 \ominus V_0$ s.t. $V_0 \oplus W_0 = V_1$. Note that generally, we have

$$W_{j-1} = V_j \ominus V_{j-1}.$$

**Lemma 4.7.** *In a Hilbert space $H$, if $w_k$ are orthogonal vectors and the sum $\sum_k \|w_k\|^2 < \infty$, then the sum $\sum_k w_k$ converges.*

By definition of direct sum, we have

$$V_n = \bigoplus_{k=-\infty}^{n-1} W_k,$$

$$V_n^{\perp} = \bigoplus_{k=n}^{\infty} W_k.$$

As a result, we have $L^2 = \bigoplus_{-\infty}^{\infty} W_n$, so we have:

**Theorem 4.8.** *Every vector $f \in L^2(-\infty, \infty)$ can be uniquely expressed as a sum $\sum_{j=-\infty}^{\infty} w_j$, where $w_j \in W_j$.*

What are the $W_j$ spaces? Consider $W_0$. Define $A$ to be the functions which are constant on half-integers and take equal and opposite values on half of each integer intervals. We claim that $W_0 = A$. In general, we can show that $W_j$ are the square integrable functions that take on equal and opposite values on each half of the dyadic intervals of length $2^{-j-1}$.

What is a basis for $W_j$? We can show that a basis for $W_j$ is $\{2^{j/2}\psi(2^j x - k)\}_{k=-\infty}^{\infty}$. Define $\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k)$. Using Theorem 4.8, we can show that every $f \in L^2$ can be written as

$$f = \sum_j \sum_k a_{jk}\psi_{jk}(x).$$

Note that the $\psi_{jk}$ form an orthonormal basis for $L^2$.

### 4.4 General Multiresolution Analysis

Suppose we use a different pixel function $\phi(x)$. Can we use this function to build approximations to general functions?

Consider the following assumptions. $|\phi(x)|$ has a finite integral s.t. $\int \phi(x)dx \neq 0$. Suppose $\phi(x)$ is a normalized function s.t. $\int \phi(x)^2 dx = 1$.

In the general construction, define $V_0$ to be all $L^2$ combinations of $\phi$ and its translates, *i.e.*, $V_0 = \{f(x) = \sum a_k\phi_{0k}(x) \mid a_k \in \mathbb{R}, f \in L^2\}$ with $\phi_{0k}(x) = \phi(x-k)$. Define $V_1, V_2, \ldots$ in a similar manner, *i.e.*,

$$V_j = \{f(x) = \sum_k a_k\phi(2^jx - k) \mid a_k \in \mathbb{R}, f \in L^2\}.$$

Note that $f(x) \in V_n \implies f(2x) \in V_{n+1}$.

We can show that the basis $\{\phi(x-k)\}$ for $V_0$ is orthogonal. Note that this is not automatic because $\phi(x-k)$ and $\phi(x-k-1)$ can overlap now. It is possible to show that $\sum |\hat{\phi}(\omega - 2n\pi)|^2 = \frac{1}{2\pi}$ is equivalent to orthonormality of $\{\phi(x-k)\}$.

What must be true of $\phi$ for $V_0 \subset V_1$ in general? We need $\phi(x)$ to be a linear combination of translates of $\sqrt{2}\phi(2x) = \sum_k h_k\phi_{1k}(x)$, where $\phi_{1k}(x) = 2^{1/2}\phi(2x-k)$.

**Example 4.9.** Suppose $\phi(x)$ is the scaling function of the Haar wavelet, then we know that

$$\begin{aligned}\phi(x) &= \phi(2x) + \phi(2x-1)\\ &= \frac{1}{\sqrt{2}}\phi_{10}(x) + \frac{1}{\sqrt{2}}\phi_{11}(x)\\ &= h_{10}\phi_{10}(x) + h_{11}\phi_{11}(x).\end{aligned}$$

Thus, in this case all $h$'s are 0 except $h_{10}$ and $h_{11}$. Note that in general that since this is an orthonormal expansion, $\sum_k h_k^2 = \|\phi(x)\|^2 < \infty$.

In general, we have

$$\phi(x) = \sum_{k=-\infty}^{\infty} h_k\phi_{1k}(x) = \lim_{N\to\infty}\sum_{k=-N}^{N} h_k\phi_{1k}(x) := \lim_{N\to\infty} F_N(x)$$

in the sense that

$$\left\|\phi(x) - \sum_{k=-N}^{N} h_k\phi_{1k}(x)\right\| \to 0.$$

**Corollary 4.10** (Corollary of Plancherel theorem)**.** *The Fourier transform is a bounded linear transformation. In particular, if the sequence of functions $\{F_N(x)\}$ converges in $L^2$ norm, then*

$$\mathcal{F}\left(\lim_{n\to\infty} F_N\right)(\omega) = \lim_{N\to\infty}\mathcal{F}(F_N)(\omega)$$

*in $L^2$ norm, i.e., Fourier transforms commute with limits.*

It follows that $\mathcal{F}\left(\sum h_k\phi_{1k}(x)\right) = \sum h_k\mathcal{F}(\phi_{1k}(\omega))$. If $\mathcal{F}(\phi)(\omega) := \hat{\phi}(\omega)$, then generally

$$\mathcal{F}(\phi_{jk})(\omega) = 2^{-j/2}e^{-i\omega k/2^j}\hat{\phi}(\omega/2^j).$$

Then we can show that

$$\hat{\phi}(\omega)(x) = \sum_{k=-\infty}^{\infty} h_k\frac{1}{\sqrt{2}}e^{-ik(\omega/2)}\hat{\phi}(\omega/2).$$

Define $m(\omega/2) = \sum h_k\frac{1}{\sqrt{2}}e^{-ik(\omega/2)}$. Note that $m$ is $2\pi$-periodic. Also note that $m(\omega) \in L^2[0, 2\pi]$, so

$$\hat{\phi}(\omega) = m(\omega/2)\hat{\phi}(\omega/2). \tag{5}$$

This condition exactly summarizes our demand that $V_0 \subset V_1$. Then it follows that $V_j \subset V_{j+1}$ in general.

Now let us introduce some preliminaries. Given a Hilbert space $H$ and a closed subspace $V$, for $f \in H$, write $f = v + v^\perp$, where $v \in V, v^\perp \in V^\perp$. The operator $P$ defined by $Pf = P(v + v^\perp) = v$ is the *orthogonal projection* onto $V$.

**Example 4.11.** Suppose $V \subset L^2[-\pi, \pi]$ is the set of even functions. Then for $f \in L^2$, $Pf(x) = f_{\text{even}}(x) = \frac{f(x)+f(-x)}{2}$.

### 4.5 General Wavelets

Now how do we construct new wavelets? Recall that the condition $\cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots$ is equivalent to $\hat{\phi}(\omega) = m_0(\omega/2)\hat{\phi}(\omega/2)$.

Also recall that the condition there exists an orthogonal basis for $V_0$ in the family of functions $\phi(x-k)$ is equivalent to $\sum |\hat{\phi}(\omega + 2\pi k)|^2 = \frac{1}{2\pi}$. If it is also the case that $\phi \in L^2(\mathbb{R})$, then $\cap V_j = \{0\}$ and $\overline{\cup V_j} = L^2(\mathbb{R})$.

**Theorem 4.12** (Conditions for an MRA)**.** *The conditions $\hat{\phi}(\omega) = m(\omega/2)\hat{\phi}(\omega/2)$ and $\sum |\hat{\phi}(\omega + 2\pi k)|^2 = \frac{1}{2\pi}$ are necessary and sufficient for the spaces $\{V_j\}$ and scaling function $\phi$ to form a multiresolution analysis.*

We can also show that $|m_0(\omega/2)|^2 + |m_0(\omega/2+\pi)|^2 = 1$. Also, since $\{\phi_{jk}(x)\}$ form a basis for $V_j$, so the wavelets $\psi_{jk}$ will form a basis for $W_j$.

Furthermore, if we consider the Fourier transform is defined as follows

$$\hat{\psi}(\omega) = e^{i\omega/2}\overline{m_0(\omega/2+\pi)}\hat{\phi}(\omega/2). \tag{6}$$

We can show that

$$\hat{\phi}(\omega) = \frac{1}{\sqrt{2\pi}} \prod_{j=1}^{\infty} m_0\left(\omega/2^j\right). \tag{7}$$

Hence, if we can find $m_0(\omega)$, then we can also find the scaling function $\phi$.

**Example 4.13.** Consider the Haar wavelet. We have $\hat{\phi}(\omega) = \frac{2}{\sqrt{2\pi}\omega} e^{-i\omega/2} \sin\frac{\omega}{2}$, so by Equation (5), we have $\hat{\phi}(\omega) = e^{-i\omega/4}\cos\frac{\omega}{2}$. Then by Equation (6), we have

$$\hat{\psi}(\omega) = -\frac{4i}{\sqrt{2\pi}\omega} e^{i\omega/2} \sin^2\frac{\omega}{4}.$$

### 4.6 Meyer Wavelets

The Meyer wavelet is defined by the scaling function

$$\hat{\phi}(\omega) = \frac{1}{\sqrt{2\pi}} \begin{cases} 1 & \text{if } |\omega| \le \frac{2\pi}{3} \\ \cos\left[\frac{\pi}{2}\nu(\frac{3}{2\pi}|\omega| - 1)\right] & \text{if } \frac{2\pi}{3} \le |\omega| \le \frac{4\pi}{3} \\ 0 & \text{otherwise,} \end{cases}$$

where $\nu$ is any infinitely differentiable non-negative function satisfying

$$\nu(x) = \begin{cases} 0 & \text{if } x \le 0 \\ 1 & \text{if } x \ge 1 \\ \text{smooth transition in } \nu : 0 \to 1 \text{ as } x : 0 \to 1, \end{cases}$$

where $\nu(x) + \nu(1-x) = 1$. We can verify that it has all the right properties in Theorem 4.12.

Note that since $\hat{\phi}(\omega)/\hat{\phi}(\omega/2) = \sqrt{2\pi}\hat{\phi}(\omega)$ in $[-2\pi, 2\pi]$, we can define $m_0(\omega/2) = \sqrt{2\pi}\hat{\phi}(\omega)$ if $\omega \in [-2\pi, 2\pi]$. The definition is ambiguous outside this interval since both the numerator and denominator is zero. Therefore, define $m_0$ by periodic extension by adding all possible translates of the bump $\hat{\phi}(\omega)$ to make it $4\pi$-periodic:

$$m_0(\omega/2) = \sqrt{2\pi} \sum_k \hat{\phi}(\omega + 4\pi k).$$

We can show that

$$\hat{\psi}(\omega) = \frac{1}{\sqrt{2\pi}} \begin{cases} e^{i\omega/2} \sin\left[\frac{\pi}{2}\nu\left(\frac{3}{2\pi}|\omega| - 1\right)\right] & |\omega| \in \left[\frac{2\pi}{3}, \frac{4\pi}{3}\right] \\ e^{i\omega/2} \cos\left[\frac{\pi}{2}\nu\left(\frac{3}{4\pi}|\omega| - 1\right)\right] & |\omega| \in \left[\frac{4\pi}{3}, \frac{8\pi}{3}\right] \\ 0 & \text{otherwise.} \end{cases}$$

**Proposition 4.14** (Properties of the Meyer wavelet)**.** *The Meyer wavelet $\psi(x)$ has the following properties:*

- *$\hat{\psi}(\omega)$ is infinitely differentiable, and one can check that all derivatives are* 0 *from both sides at the break.*

- *The support of $\hat{\psi}(\omega)$ is on a finite interval.*
- *$\psi(x)$ decays at $\infty$ faster than any inverse power of $x$.*
- *$\psi(x)$ is infinitely differentiable.*

**Claim 4.15.** *$\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k)$ forms an orthonormal basis for $L^2(\mathbb{R})$.*

### 4.7 Daubechies Wavelets

**Theorem 4.16** (Cohen, 1990)**.** *If the trigonometric polynomial $m_0$ satisfies $m_0(0) = 1$, $|m_0(\omega)|^2 + |m_0(\omega + \pi)|^2 = 1$, and $m_0(\omega) \neq 0$ for $|\omega| \leq \pi/3$, then $\sum |\hat{\phi}(\omega + 2\pi k)|^2 = \frac{1}{2\pi}$ is satisfied by Equation (7).*

A choice of $m_0$ is

$$m_0(\omega) = \frac{1}{8}[(1+\sqrt{3}) + (3+\sqrt{3})e^{-i\omega} + (3-\sqrt{3})e^{-2i\omega} + (1-\sqrt{3})e^{-3i\omega}].$$

Recall that once we have an orthonormal wavelet basis $\{\psi_{jk}\}$, we can write any function

$$f(x) = \sum_{j,k} a_{jk}\psi_{jk}(x),$$

where $a_{jk} = \langle f, \psi_{jk}\rangle$. Numerically, we can find $a_{jk}$ using numerical integration on the computer. There are very efficient methods of doing this: for one wavelet, all others are rescalings and translations of the original one.

### 4.8 General Properties of Orthonormal Wavelet Bases

**Theorem 4.17.** *If the basic wavelet $\psi(x)$ has exponential decay, then $\psi$ cannot be infinitely differentiable.*

Recall that the Haar wavelets had compact support. When will wavelets have compact support in general?

For $\phi(x) \in V_0 \subset V_1$, we have for some choice of $h_k$:

$$\phi(x) = \sum_k h_k \sqrt{2}\phi(2x - k).$$

The constants $h_k$ relate $V_0$ to $V_1$.

**Theorem 4.18.** *Suppose the $h_k$ are defined as above. $\psi, \phi$ have compact support iff finitely many $h_k \neq 0$.*

Suppose we start with a finite sequence of numbers $h_k$. Construct

$$m_0(\omega) = \sum_k \frac{h_k}{\sqrt{2}} e^{-ik\omega},$$

$$\hat{\phi}(\omega) = \frac{1}{\sqrt{2\pi}} \prod_{j=1}^{\infty} m_0(2^{-j}\omega),$$

$$\hat{\psi}(\omega) = e^{i\omega/2}\overline{m_0(\omega/2 + \pi)}\hat{\phi}(\omega/2).$$

Finally, take the inverse Fourier transform to get the wavelet $\psi(x)$.

Do wavelet expansions actually converge to the function being expanded at individual points?

**Theorem 4.19.** *Assume that scaling function $\phi$ is bounded by an integrable decreasing function. If $f \in L^2$, then the wavelet expansion of $f$ converges pointwise a.e. to $f$.*

How fast do wavelet expansions converge to $f$? It depends on how "regular" the wavelet is.

**Theorem 4.20.** *In $d$ dimensions, the wavelet expansion of $f$ converges to a smooth $f$ in such a way that the partial sum*

$$\sum_{j \leq N, k} a_{jk}\psi_{jk}(x)$$

*differs from $f(x)$ at each $x$ by at most $C \cdot 2^{-Ns}$ iff*

$$\int |\hat{\psi}(\omega)|^2 |\omega|^{-2s-d}\, d\omega < \infty.$$

### 4.9 Continuous Wavelet Transforms

Consider a function $\psi(x) \in L^2$ s.t. $\psi$ decays faster than $1/x^2$ at $\infty$ and $\int \psi(x)dx = 0$. Then we can define an integral wavelet expansion using rescalings of $\psi(x)$ with rescaled functions

$$\psi_{a,b} = |a|^{1/2}\psi(a(x-b)),$$

where $a, b \in \mathbb{R}$. Thus, $a$ is a dilation parameter and $b$ is a translation parameter. Now define the wavelet transform as follows:

$$(Wf)(a,b) = \langle \psi_{a,b}, f \rangle = \int \overline{\psi_{a,b}(x)}\, f(x)\, dx.$$

How do we recover $f$ from $(Wf)(a,b)$?

**Claim 4.21.** *We have*

$$f(x) = C \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (Wf)(a,b)\psi_{a,b}\, da\, db,$$

*where*

$$C^{-1} = -2\pi \int |\omega|^{-1} |\hat{\phi}(\omega)|^2\, d\omega.$$

### 4.10 Application of the Integral Wavelet Transform: Image Reconstruction

Dyadic wavelet transforms are a variation on the continuous wavelet transform. In this case, we only allow dilations by powers of 2:

$$\psi_{j,b}(x) = 2^j \psi(2^j(x-b)).$$

Define $\psi_j(x) = 2^j\psi(2^j x)$. Then this dyadic wavelet transform is defined by

$$\begin{aligned}(Wf)(j,b) &= \int f(x)\psi_{j,b}(x)\, dx.\\ &= (f * \psi_j)(b).\end{aligned}$$

Assume that the Fourier transform $\hat{\psi}(\omega)$ satisfies $\sum_{-\infty}^{\infty} |\hat{\psi}(2^j\omega)|^2 = \frac{1}{2\pi}$. Then under these assumptions, we can show that we can recover $f$ in this casel. The recovery formula for $f$ is

$$f(x) = \sum_{j=-\infty}^{\infty} (Wf)(j,x) * \psi_j(-x).$$

Given $f(x)$, what sort of function is the wavelet transform $(Wf)(j,b)$ as a function of $j$ and $b$?

Suppose $V$ is the collection of possible wavelets. When is an arbitrary function $g(j,b)$ a wavelet transform? We can check that $g$ must satisfy a so-called reproducing kernel equation. $g(j,b)$ is the wavelet transform of some function iff

$$g(j,b) = (Kg)(j,b) := \sum_{\ell=-\infty}^{\infty} \psi_j(b) * \psi_\ell(-b) * g(\ell, b).$$

Let us consider recovering $f$ from the wavelet transform. We can recover $f$ as a sum of $f$ at different scales:

$$f = \sum_{j=-\infty}^{\infty} (Wf)(j,x) * \psi_j(-x).$$

Since $\psi$ is a known function, we can recover $f$ from a sequence of functions. Assume $a(x)$ is a cubic B-spline. Then we let the wavelet be its first derivative: $\psi(x) = a'(x)$.

**Conjecture 4.22.** *We can recover $f$ not from knowing all of the functions $W(j,x)$, but just from knowing their maxima and minima.*

Meyer proved this conjecture false for certain choices of $\psi$. However, it is true for the derivative of the Gaussian: $\psi(x) = [e^{-x^2}]'$.

Assume that we are given only the maxima and minima points of the function $W(j,x)$ for each $j$. How do we recover $f$?

Define $\Gamma$ to be the set of all functions $g(j,x)$ which have the same set of maxima and minima (in $x$) as $W(j,x)$ for each $j$ and $V$ to be the set of all $g(j,x)$ which are wavelet transforms of some function of $x$. The idea is that $Wf \in \Gamma \cap V$.

The algorithm is as follows:

1. Start with only the maxima information about $Wf(j,x)$. Call $M$ the maxima information.
2. Make an initial guess of the full $Wf(j,x)$ using any function $g_1(j,x)$ which has the same maxima as $Wf(j,x)$.
3. Find the closest function $g_2(j,x) \in V$ to $g_1(x)$.
4. Find the closest function $g_3(j,x) \in \Gamma$ to $g_2$.
5. Continue this way: at each stage $j$: find the closest function $g_j$ to $g_{j-1}$ in the space $V$ or $\Gamma$ (alternatingly).
6. Eventually $g_j(j,x) \xrightarrow{j\to\infty} Wf(j,x)$.

The conclusion is that we can recover the wavelet transform of a function just by knowing its maxima in $x$. As a result, we can use this for compression of images.

### 4.11 Wavelets and Wavelet Transforms in Two Dimensions

MRA and wavelets can be generalized to higher dimensions. The usual choice for a two-dimensional scaling function or wavelet is a product of two one-dimensional functions.

For example, $\phi_2(x,y) = \phi(x)\phi(y)$ and scaling equation has form $\phi(x,y) = \sum_{k,l} h_{kl} \cdot 2\phi(2x-k, 2y-l)$. Since both $\phi(x), \phi(y)$ satisfy the scaling equation, we have $h_{kl} = h_k h_l$. Thus the two dimensional scaling equation is the product of two one-dimensional scaling equations. Hence, we can proceed analogously to construct wavelets using one-dimensional functions. However, now there are three types of basic wavelets:

$$\psi^{(I)}(x,y) = \phi(x)\psi(y)$$
$$\psi^{(II)}(x,y) = \psi(x)\phi(y)$$
$$\psi^{(III)}(x,y) = \psi(x)\psi(y).$$

## 5 Topological Spaces

A *topological space* is a set $S$ together with a distinguished family $\mathcal{F}$ of subsets of $S$ called *open sets* with the following properties:

(i) $A, B \in \mathcal{F} \implies A \cap B \in \mathcal{F}$.

(ii) $A_\beta \in \mathcal{F}$ for $\beta \in B \implies \cup_{\beta \in B} A_\beta \in \mathcal{F}$.

(iii) $\varnothing, S \in \mathcal{F}$.

If $x \in S$, a set $N$ is a *neighborhood* of $x$ if there exists an open set $U$ s.t. $x \in U \subset N$. A set $F \subset S$ is *closed* if $F^c$ is open.

Let $S, T$ be two topological spaces. A function $f : S \to T$ is *continuous* if $f^{-1}(U)$ is open for any open set $U \subset T$.

A set $A \subset S$ is *dense* if every open set in $S$ contains at least one point in $A$. $S$ is *separable* if it contains a countable subset $A$ which is dense in $S$.

A family $\mathcal{N}$ of subsets of $S$ is a *neighborhood base* at $x \in S$ if each $N \in \mathcal{N}$ is a neighborhood of $x$ and for any other neighborhood $M$ containing $x$, $\exists N \in \mathcal{N}$ s.t. $N \subset M$.

A family $\mathcal{N}$ of subsets of $S$ is a *base* for $S$ if every open set $M \subset S$ can be expressed as a union of sets in $\mathcal{N}$, *i.e.*,

$$M = \bigcup_{A \in \mathcal{N} : A \subset M} A.$$

$S$ is *first countable* if each $x \in S$ has a countable neighborhood base. $S$ is *second countable* if $S$ has a countable base.

We have the following:

(a) A topological space is called a $T_1$ *space* iff $\forall x, y$ s.t. $x \neq y$, there exists an open set $O$ with $y \in O, x \notin O$. Equivalently, a space is $T_1$ iff $\{x\}$ is closed $\forall x$.

(b) A topological space is called *Hausdorff* (or $T_2$) if $\forall x, y$, there exists disjoint open sets $O_1, O_2$ s.t. $x \in O_1, y \in O_2$.

(c) A topological space is called *regular* (or $T_3$) iff it is $T_1$ and $\forall x$ and $C$ closed with $x \notin C$, there exists open sets $O_1, O_2$ s.t. $x_1 \in O_1, C \subset O_2, O_1 \cap O_2 = \varnothing$. Equivalently, a space is $T_3$ if the closed neighborhoods of any point are a neighborhood base.

(d) A topological space is called *normal* (or $T_4$) iff it is $T_1$ and $\forall C_1, C_2$ closed with $C_1 \cap C_2 = \varnothing$, there exists open sets $O_1, O_2$ with $C_1 \subset O_1, C_2 \subset O_2, O_1 \cap O_2 = \varnothing$.

**Proposition 5.1.** $T_4 \implies T_3 \implies T_2 \implies T_1$.

**Proposition 5.2.** *We have the following:*

*(i) Every metric space is first countable.*

*(ii) A metric space is second countable iff it is separable.*

*(iii) Any second countable topological space is separable.*

A topological space is called *disconnected* iff it contains a nonempty proper subset which is both open and closed. Equivalently, a topological space is disconnected iff it can be written as the union of two disjoint nonempty closed sets.

### 5.1 Nets

Nets are generalizations of sequences appropriate to general topological spaces, not just metric spaces.

**Example 5.3.** Consider the limit $\lim_{n\to\infty} \frac{n}{n+1}$ where $n \in \mathbb{N}$. This is a limit of a sequence. Then $\lim_{x\to\infty} \frac{x}{x+1}$ where $x \in \mathbb{R}$ is a limit of a net.

A *directed system* is a set $I$ together with an order $\prec$ on $I$ which satisfies

(a) if $\alpha, \beta \in I, \exists \gamma \in I$ s.t. $\alpha \prec \gamma, \beta \prec \gamma$.

(b) $\prec$ is a partial ordering.

A *net* in $S$ is a set $\{x_\alpha\}_{\alpha\in I} \subset S$, where $I$ is a directed sequence. If $I = \mathbb{N}$, then the net is called a *sequence*.

**Example 5.4.** $I$ can be the collection of all subsets of a set $A$. If $\alpha, \beta \in I$, then $\alpha \prec \beta$ if $\alpha \subset \beta$.

Define $x_\alpha = \frac{1}{|\alpha|} \in S = \mathbb{R}$. Then $\{x_\alpha\}$ is a net, and $I$ is a directed system. A net $\{x_\alpha\}$ *converges* to a point $x \in S$ if for any neighborhood $N$ of $x$, $\exists \beta$ s.t. $x_\alpha \in N$ for $\alpha \succ \beta$.

**Theorem 5.5.** *Let $A \subset S$, where $S$ is a topological space. Then a point $x \in \overline{A}$ iff there exists a net $\{x_\alpha\}_{\alpha\in I} \subset A$ s.t. $x_\alpha \to x$.*

**Theorem 5.6.** *If $S, T$ are topological spaces. $f : S \to T$ is continuous iff for every net $\{x_\alpha\} \subset S$ which converges, the net $\{f(x_\alpha)\} \subset T$ also converges.*

**Theorem 5.7.** *In a Hausdorff space, any net $\{x_\alpha\}$ can have at most one limit.*

A net $\{y_\beta\}_{\beta\in J}$ is a *subnet* of the net $\{x_\alpha\}_{\alpha\in I}$ if

(a) $\{y_\beta\}_{\beta\in J} \subset \{x_\alpha\}_{\alpha\in I}$.

(b) $\forall \alpha \in I, \exists \beta \in J$ s.t. if $\beta' \succ \beta$, then $y_{\beta'} = x_{\alpha'}$ for an $\alpha' \succ \alpha$.

**Proposition 5.8.** *If a net $x_\alpha \to x$, then every subnet $y_\beta$ of $x_\alpha$ also converges to $x$.*

Let $(S, \mathcal{F})$ be a topological space. Let $A \subset S$. Define $\mathcal{F}_A = \{O \cap A \mid O \in \mathcal{F}\}$. Then $(A, \mathcal{F}_A)$ is a topological space, and $\mathcal{F}_A$ is called the *relative topology* on $A$.

### 5.2 Compactness

Given a collection of sets $\mathcal{C}$, we say that $\mathcal{C}$ is a *cover* of (or covers) another set $F$ if $F$ is contained in the union of the sets in $\mathcal{C}$. We say that $\mathcal{C}$ is an *open cover* of $F$ if all sets in $\mathcal{C}$ are open. A set $A \subset S$ is *compact* if any open cover of $A$ has a finite subcover.

**Theorem 5.9** (Heine-Borel theorem)**.** *$S \subset \mathbb{R}^n$ is compact iff it is closed and bounded.*

**Theorem 5.10** (Bolzano-Weierstrass theorem)**.** *A set $A$ is compact iff every net in $A$ has a convergent subnet.*

**Proposition 5.11.** *We have the following:*

*(a) A closed subset of a compact set is compact.*

(b) *Suppose $S, T$ are topological spaces. If $f : S \to T$ is continuous and $A \subset S$ is compact, then $f(A)$ is compact.*

**Proposition 5.12.** *In a Hausdorff space, every compact set is closed.*

Let $S, T$ be topological spaces. Let $f : S \to T$ be a continuous bijection with $f^{-1}$ continuous. Then, $f$ is a *homeomorphism.*

**Theorem 5.13.** *If $S, T$ are compact Hausdorff spaces, and $f : S \to T$ is continuous and a bijection, then $f$ is a homeomorphism.*

A topological space $S$ is *normal* if

(i) $\forall x, y \in S$, there exists an open set $O \subset S$ s.t. $x \in O, y \notin O$.

(ii) If $C_1, C_2 \subset S$ where $C_1, C_2$ are closed and disjoint, then there exists disjoint open sets $O_1, O_2$ s.t. $C_1 \subset O_1, C_2 \subset O_2$.

**Theorem 5.14** (Urysohn's lemma)**.** *Let $X$ be a normal space, and $C_1, C_2$ be disjoint closed sets in $X$. Then there exists a continuous function $f : X \to \mathbb{R}$ with $0 \leq f(x) \leq 1$ s.t. $f(x) = 0$ if $x \in C_1$ and $f(x) = 1$ if $x \in C_2$.*

**Lemma 5.15.** *A continuous function on a compact set $K$ is bounded.*

**Theorem 5.16.** *Let $X$ be a compact Hausdorff space. Let $C(X) = \{f : X \to \mathbb{C} \mid f \text{ is continuous}\}$ with norm $\|f\|_\infty = \sup |f(x)|$. Then $C(X)$ is a Banach space with this norm.*

Define $C_{\mathbb{R}}(X)$ to be the same of $C(X)$, using only real-valued functions.

A collection $\mathcal{G} \subset C(X)$ is a *subalgebra* of $C(X)$ if

(i) $\mathcal{G}$ is a subspace of $C(X)$, *i.e.*, it is closed under addition and scalar multiplication.

(ii) $\mathcal{G}$ is closed under multiplication, *i.e.*, if $g_1, g_2 \in \mathcal{G}$, then $g_1 g_2 \in \mathcal{G}$.

$\mathcal{G}$ *separates points* in $X$ if $\forall x_1, x_2 \in X, \exists g \in \mathcal{G}$ s.t. $g(x_1) \neq g(x_2)$.

**Theorem 5.17** (Stone-Weierstrass theorem)**.** *Let $X$ be a compact Hausdorff space. Let $\mathcal{G}$ be a subalgebra of $C_{\mathbb{R}}(X)$ which is closed as a subset of $C_{\mathbb{R}}(X)$, and let $\mathcal{G}$ separate points in $X$. Then either $\mathcal{G} = C_{\mathbb{R}}(X)$ or $\mathcal{G} = \{g \in C_{\mathbb{R}}(X) \mid g(x_0) = 0\}$ for some fixed $x_0$. Thus, if $g(x) = 1 \in \mathcal{G}$, then $\mathcal{G} = C_{\mathbb{R}}(X)$.*

**Corollary 5.18.** *Real-valued polynomials $\mathcal{P}$ are dense in $C_{\mathbb{R}}[a, b]$.*

**Remark 5.19.** If we replace $C_{\mathbb{R}}(X)$ by $C(X)$, for the Stone-Weierstrass theorem to be true, we require $\mathcal{G}$ to be a *star-algebra*, *i.e.*, it is closed under the operation $f \to f^*$ (complex conjugation of functions).

### 5.3 Measure Theory on Compact Spaces

Suppose $X$ is a compact Hausdorff space. A set $A$ in a topological space is a $G_\delta$ set if it is the intersection of a countable number of open sets.

**Proposition 5.20.** *If $X$ is a compact Hausdorff space, and $f \in C_{\mathbb{R}}(X)$, then $f^{-1}([a, \infty))$ is a compact $G_\delta$ set.*

The $\sigma$-algebra generated by the compact $G_\delta$ sets is called the *Baire sets*. Functions measurable w.r.t. the Baire sets are the *Baire functions*. A measure on the Baire sets is a *Baire measure* if it is finite.

**Remark 5.21.** Note that the Baire sets is a subset of Borel sets since Borel sets contain all compact $G_\delta$ sets. In $\mathbb{R}$, the Borel sets are the Baire sets since Borel sets are generated by closed intervals $[a, b]$, which are compact $G_\delta$ sets.

**Theorem 5.22.** *Every Baire measure is regular, i.e., if $A$ is a Baire set, then $\mu(A) = \inf\{\mu(O) \mid O \text{ open, Baire, } O \supset A\}$ and $\mu(A) = \sup\{\mu(K) \mid K \text{ compact, Baire, } K \subset A\}$.*

**Theorem 5.23.** *We have the following:*

(a) *If $\mu$ is a measure on the Baire sets, then $\mu$ has a unique extension to a regular measure on the Borel sets.*

(b) *If $\mu$ is a Baire measure which has been extended to a regular Borel measure, then every Borel set differs by at most a set of measure 0 for some Baire set.*

We continue to assume that $C(X)$ and the $C_{\mathbb{R}}(X)$ have the usual supremum norm. The theorems below apply to both $C(X)$ and $C_{\mathbb{R}}(X)$.

Let $\ell : C(X) \to \mathbb{C}$ (or $C_{\mathbb{R}}(X) \to \mathbb{R}$) be a continuous linear functional with $\ell(f) \geq 0 \;\forall f$ with $f(x) \geq 0 \;\forall x \in X$. Then $f$ is called a *positive* linear functional.

**Proposition 5.24.** *If $\ell$ is a positive linear functional, then $\|\ell\| = \ell(1)$.*

**Example 5.25.** Let $X$ be a compact Hausdorff space and $\mu$ be a Baire measure. $\forall f \in C(X)$, define $\ell(f) = \int_X f(x)\, d\mu$. Then $\ell$ is a positive linear functional.

**Theorem 5.26** (Riesz-Markov theorem). *Let $X$ be a compact Hausdorff space, and let $\ell$ be a positive linear functional on $C(X)$. Then $\exists!$ Baire measure $\mu$ on $X$ s.t.*

$$\ell(f) = \int_X f\, d\mu \;\forall x \in C(X).$$

What is the dual of $C_{\mathbb{R}}(X)$?

**Theorem 5.27.** *Let $\ell$ be a bounded linear functional on $C_{\mathbb{R}}(X)$, i.e., $\ell \in C_{\mathbb{R}}(X)^*$. Then $\ell = \ell_+ - \ell_-$, where $\ell_+, \ell_-$ are unique positive linear functionals. Also, $\|\ell\| = \|\ell_+\| + \|\ell_-\| = \ell_+(1) + \ell_-(1)$.*

**Theorem 5.28** (General Riesz-Markov theorem). *If $\ell$ is a general linear functional on $C_{\mathbb{R}}(X)$, then*

$$\ell(f) = \int_X f(x)\, d\mu,$$

*where $\mu$ is a signed measure, i.e., a generally nonpositive measure.*

A topological space $X$ is *locally compact* if every $p \in X$ has a compact neighborhood.

**Remark 5.29.** The Riesz-Markov theorem can be extended from compact spaces to locally compact spaces, *e.g.*, extending from $[0, 1]$ to $\mathbb{R}$.

Recall that if $\mathcal{T}_1, \mathcal{T}_2$ are two topologies, then $\mathcal{T}_1$ is *weaker* than $\mathcal{T}_2$ if $\mathcal{T}_1 \subset \mathcal{T}_2$. If $\mathcal{G}$ is a collection of sets in $X$, $\mathcal{T}$ is the topology *generated* by $\mathcal{G}$ if $\mathcal{T}$ is the weakest topology where all sets in $\mathcal{G}$ are open. We can show that $\mathcal{T}$ is all possible unions of finite intersections of sets in $\mathcal{G}$.

Let $X$ be a Banach space and $X^*$ be its dual space. The *weak topology* on $X$ is the weakest topology on $X$ s.t. every linear functional $\ell \in X^*$ is continuous on $X$.

Suppose $\mathcal{T}$ is this topology. What are the open sets? Notice that $\mathcal{T}$ is the topology generated by all sets of the form $\ell^{-1}(O)$, where $O \subset \mathbb{R}$ is open and $\ell \in X^*$. We can show that $\mathcal{T}$ is all possible unions of sets of the form

$$\bigcap_{i=1}^{n} \ell_i^{-1}(I_i) := N(\ell_1, \ldots, \ell_n; I_1, \ldots, I_n),$$

where $\ell_i \in X^*, I_i \subset \mathbb{R}$ is an open interval. All sets of this form make a base for the topology $\mathcal{T}$. Additionally, sets of the form

$$N(\ell_1, \ldots, \ell_n; \varepsilon)$$

form a *neighborhood base* at $x = 0$.

**Theorem 5.30.** *In the weak topology $\mathcal{T}$, the net $x_\alpha \to x$ iff $\ell(x_\alpha) \to \ell(x) \;\forall \ell \in X^*$.*

**Proposition 5.31.** *We have the following:*

*(a) The weak topology is weaker (not necessarily strictly weaker) than the norm topology (usual, strong topology) on $X$.*

*(b) If $x_\alpha \to x$ in the weak topology, then $\|x_\alpha\|$ are all bounded.*

*(c) The weak topology is Hausdorff.*

**Example 5.32.** Let $X$ be a compact Hausdorff space. Note that $C(X)$ is a Banach space. Then, if $f_n$ is a sequence of functionals in $C(X)$, we have $f_n \to f$ weakly iff $f_n(x) \to f(x) \;\forall x$ and $\|f_n\|$ are all bounded uniformly.

### 5.4 The Weak* Topology

Let $X$ be a Banach space and $X^*$ be its dual space. The weak* topology is the weakest topology on $X^*$ for which all functions $L_x, x \in X$ are continuous, where $L_x : X^* \to \mathbb{R}$ is defined by $L_x(\ell) = \ell(x)$.

**Example 5.33.** Let $X$ be a compact Hausdorff space. Define $\mathcal{M}(X) = C_{\mathbb{R}}(X)^*$. We claim that $\mathcal{M}(X) \cong$ set of all signed finite Baire measures on $X$ via the correspondence $\ell \longleftrightarrow \mu$, where $\ell \in \mathcal{M}(X)$ and $\mu$ is a signed measure on $X$ s.t.

$$\ell(f) = \int_X f(x)\, d\mu.$$

Consider the weak* topology on $C(X)^*$. This is called the *vague* topology on $\mathcal{M}$, which arises in probability. If $X = \mathbb{R}$, and $\mu_n, \mu \in \mathcal{M}$, then the notion of $\mu_n \to \mu$ in the vague topology is the same as the notion of *weak convergence* in probability theorem. For $x \in X$, define $\delta_x$ to be a measure on $X$ s.t.

$$\delta(A) = \begin{cases} 0 & \text{if } x \notin A, \\ 1 & \text{if } x \in A. \end{cases}$$

**Claim 5.34.** *Measures of the form*

$$\sum_{i=1}^{n} c_i \delta_{x_i},$$

*where $c_i \in \mathbb{R}, x_i \in X$ are weak*-dense in the set of all signed Baire measures.*

**Remark 5.35.** [Riesz-Markov correspondence] For a fixed $x_i$, there is a one-to-one correspondence between Baire measure $\delta_{x_i}$ and linear functional $\ell_{x_i}$, where $\ell_{x_i}$ is defined by $\ell_{x_i}(f) = f(x_i)$. The correspondence is given by

$$\ell_{x_i}(f) = \int f\, d\delta_{x_i}.$$

Thus, any Baire measure $\mu$ on $\mathbb{R}$ can be approximated by linear combinations of point mass measures $\delta_{x_i}$, *i.e.*,

$$\mu \approx \sum_{i=1}^{n} c_i \delta_{x_i}.$$

Correspondingly, for a continuous function $f$,

$$f(x)\, d\mu \approx \int f(x)\, d\left(\sum_{i=1}^{n} c_i \delta x_i\right) = \sum_{i=1}^{n} c_i f(x_i).$$

## 6 Ergodic Theory

Ergodic theory is about classical mechanics and dynamical systems. Suppose we have a phase space $\Omega$ of all possible states of a physical system, where $\Omega$ can be a metric or topological space.

For $t \geq 0, x \in \Omega$, define $T_t x$ to be the state of the system at state $x \in \Omega$ after time $t$, where $T_t : \Omega \to \Omega$. Assume $T_t$ is invertible $\forall t$, and $T_t(x)$ is measurable in variables $(t, x)$ jointly on $\mathbb{R} \times \Omega$ as a measure space using Borel measurable sets on $\mathbb{R}$ and $\Omega$, respectively.

Let $E \geq 0$. We usually consider the subset $\Omega_E \subset \Omega$ consisting of all states in $\Omega$ with energy $E$. The system initially in $\Omega_E$ stays in $\Omega_E$ due to energy conservation. Suppose the initial state is $x$, and at time $t$, the state is $T_t x$. Does $T_t x$ have a limit as $t \to \infty$? Not usually. However, there is usually a density that the orbit has as it moves around $\Omega_E$. Henceforth, WLOG, assume $\Omega_E$ is all of $\Omega$ since we are restricted in any case to $\Omega_E$ $\forall t$.

Assume $\Omega$ is a compact, Hausdorff topological space. Let $f$ be a continuous function on $\Omega$. Does $f(T_t x)$ have a limiting density of values? Notice that the average value from time 0 to time $\tau$ is

$$\frac{1}{\tau} \int_0^\tau f(T_t x)\, dt.$$

Does this have a limit as $\tau \to \infty$ that is independent of $x$? If yes, this is called the zeroth law of thermodynamics. If this limit exists and is independent of $x$, define

$$\ell(f) = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau f(T_t x)\, dt.$$

Note that we can easily check that $\ell(f)$ is a positive linear functional on $C(\Omega)$ and is bounded. Thus, by the Riesz-Markov theorem, there exists a Baire measure $\mu$ s.t. $\ell(f) = \int_\Omega f(x')\,d\mu(x')$. Thus, we have

$$\lim_{\tau\to\infty} \frac{1}{\tau} f(T_t x)\,dt = \int_\Omega f(x')\,d\mu(x').$$

Note that $\{T_t\}$ is a *one parameter group* of operators, *i.e.*, that it depends on a single parameter $t$, and that $T_t T_s x = T_{t+s} x$. This follows just by definition of $T_t$, since the effect of time $s$ followed by effect of time $t$ is the effect of time $s+t$. Then we can show that

$$\int_\Omega f(x)\,d\mu(x) = \int_\Omega f(x)\,d(\mu\circ T_s^{-1})(x),$$

where $\mu\circ T^{-1}$ is a new measure defined by $(\mu\circ T_s^{-1})(F) := \mu(T_s^{-1}F)$, so $\mu\circ T_s^{-1} = \mu$, *i.e.*, $T_s$ is *measure preserving* w.r.t. the measure $\mu$. Equivalently, $\mu$ is an invariant measure for $T$. Hence, if the limit $\lim_{\tau\to\infty} \frac{1}{\tau}\int_0^\tau f(T_t x)\,dt$ exists, then $T$ has an invariant measure $\mu$.

Now consider the converse question. If $T_t$ has an invariant measure $\mu$, does the above limit exist? We want to prove that $\ell(f)$ exists for $f \in C(\Omega)$. Define a new map $U_t$ on $L^2(\Omega)$ by

$$U_t f(x) = f(T_t x).$$

**Lemma 6.1** (Koopman's lemma)**.** *$U_t$ is a unitary map on $L^2(\Omega)$, i.e., it preserves $L^2$ norms.*

Note that $U_{t+s} = U_t U_s$ again. Thus, if $n\in\mathbb{Z}$, then $U_n = U_{1+\cdots+1} = U_1^n$. Now replace the above question by a slightly simpler discrete problem by setting $\tau = N \in \mathbb{Z}$. Then,

$$\begin{aligned}\frac{1}{\tau}\int_0^\tau U_t f(x)\,dt &= \frac{1}{N}\int_0^\tau U_t f(x)\,dt\\ &\approx \frac{1}{N}\sum_{m=0}^{N-1} U_m f(x).\end{aligned}$$

Does this limit as $n\to\infty$ exist?

For a linear operator $A : \mathcal{H}\to\mathcal{H}$, we define

$$\begin{aligned}\operatorname{ran} A &= \{Ax \mid x\in\mathcal{H}\},\\ \ker A &= \{x \mid Ax = 0\}.\end{aligned}$$

Given a linear bounded operator $A$ in a Hilbert space $\mathcal{H}$, the *adjoint operator* $A^*$ is uniquely defined by $\langle x, Ay\rangle = \langle A^* x, y\rangle$.

**Lemma 6.2.** *We have the following:*

*(a) If $U$ is unitary, then $Uf = f$ iff $U^* f = f$.*

*(b) For any operator $A$, $(\operatorname{ran} A)^\perp = \ker A^*$.*

**Remark 6.3.** For a real finite-dimensional v.s. (as in linear algebra), statement (b) becomes $(\operatorname{ran} A)^\perp = \ker A^\top$.

**Theorem 6.4** (Mean ergodic theorem, discrete version)**.** *Let $U$ be a unitary operator on $\mathcal{H}$. Let $P$ be the orthogonal projection onto $J = \{\psi \mid U\psi = \psi\}$, which implies $P^2 = P$ and $P(J) = J$. Then $\forall f\in\mathcal{H}$,*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{m=0}^{N-1} U_m f$$

*exists and equals $Pf$.*

**Theorem 6.5** (Mean ergodic theorem, continuous version)**.** *Let $T_t(x)$ be a one parameter group of measure preserving transformations of $(\Omega,\mu)$. Then for $f\in L^2(\Omega)$,*

$$\lim_{\tau\to\infty}\frac{1}{\tau}\int_0^\tau f(T_t w)\,dt = \int_0^1 (Pf)(T_t w)\,dt,$$

*where $P$ is the orthogonal projection onto the subspace $\{\psi \mid U_1\psi = \psi\}$, and $U_1$ is the time one map $U_1 f = f(T_1 w)$.*

Let $\mu$ is a measure on set $\Omega$ and $T : \Omega \to \Omega$ be a transformation that preserves $\mu$. We assume that $\mu$ is a probability measure for now, *i.e.*, $\mu(\Omega) = 1$.

For a single $t$, the map $T_t$ is *ergodic* if $f \circ T_t = f$ implies that $f$ is constant. The *family* $\{T_t\}_{t\geq 0}$ is ergodic if $T_t$ is ergodic $\forall t$. This definition implies that $T_1$ mixes sets up.

**Claim 6.6.** *This means $T_1$ leaves no set $A$ invariant if $A$ is nontrivial, i.e., not of full or 0 measure. That is, we never have $T_1 A = A$.*

**Proposition 6.7.** *$T_t$ is ergodic (for a fixed $t$) iff the only sets left invariant by $T_t$ have measure 0 or measure 1.*

**Corollary 6.8** (of the mean ergodic theorem)**.** *Let $T_t$ be ergodic and measure preserving on $(\Omega, \mu)$. Then $\forall f \in L^2(\Omega, d\mu)$,*

$$\lim_{\tau\to\infty} \frac{1}{\tau}\int_0^\tau f(T_t w)\,dt = \int_\Omega f(x)\,d\mu(x),$$

*i.e., the time average is equal to the space average.*

**Theorem 6.9** (Birkhoff ergodic theorem)**.** *Let $T$ be a measure preserving transformation on a measure space $(\Omega, \mu)$ (not necessarily finite). Then $\forall f \in L^1(\Omega, \mu)$, the limit*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) = f^\#(x) \in L^1(\mu)$$

*exists pointwise a.e. Further, $f^\#(x) = f^\#(Tx)$ $\forall x \in \Omega$. If $\mu(\Omega) < \infty$, then*

$$\int f^\#(x)\,d\mu = \int f(x)\,d\mu.$$

*Furthermore, if $T$ is ergodic and $\mu(\Omega) = 1$, then*

$$f^\#(x) = \int f(y)\,d\mu(y)$$

*for almost all $x$.*

# 7 Spaces of Bounded Operators

Let $X, Y$ be Banach spaces. Let $\mathcal{L}(X, Y)$ be the set of all continuous linear maps from $X$ to $Y$. Recall that if $T \in \mathcal{L}(X, Y)$, $\|T\| = \sup_{x\in X} \|Tx\|/\|x\|$, so $\mathcal{L}(X, Y)$ is a Banach space with this norm.

## 7.1 Topologies on Operator Spaces

Now let us consider topologies on $\mathcal{L}(X, Y)$. The normal topology on $\mathcal{L}(X, Y)$ is called the *uniform operator topology* (or the norm topology). Here, $\{T_n\} \to T$ iff $\|T_n - T\| \to 0$.

There are two other topologies on $\mathcal{L}(X, Y)$. Let $x \in X$. Consider the map $E_x : \mathcal{L}(X, Y) \to Y$ defined by $E_x(T) = Tx$. The *strong operator topology* is the weakest topology on $\mathcal{L}(X, Y)$ s.t. all maps $E_x$ are continuous from $\mathcal{L}(X, Y)$ to $X$ in the norm topology on $\mathcal{L}(X, Y)$.

**Remark 7.1.** In the strong operator topology, the strongly convergent nets $\{T_\alpha\} \subset \mathcal{L}(X, Y)$ are those where $\{T_\alpha x\}$ converges $\forall x \in X$.

Let $x \in X, \ell \in Y^*$. Consider the map $E_{x,\ell} : \mathcal{L}(X, Y) \to \mathbb{C}$ defined by $E_{x,\ell}(T) = \ell(Tx) \in \mathbb{C}$. The *weak operator topology* on $\mathcal{L}(X, Y)$ is the weakest topology s.t. all maps of this form are continuous.

**Remark 7.2.** In the weak operator topology, the convergent nets $\{T_\alpha\}$ are those where $\ell(T_\alpha x)$ converges $\forall x \in X, \ell \in Y^*$.

**Claim 7.3.** *Weakest topology $\subset$ strong topology $\subset$ norm topology.*

**Example 7.4.** Suppose $\ell^2$ are all sequences $\{a_n\}$ s.t. $(\sum a_n^2)^{1/2} = \|\{a_n\}\|_2 < \infty$.

(a) Consider the operator $T_n(a_1, a_2, \dots) = \frac{1}{n}(a_1, a_2, \dots)$. Then

$$\|T_n\| = \sup_{a\in\ell^2} \frac{\|T_n a\|}{\|a\|} = \sup_{a\in\ell^2} \frac{\|a/n\|}{\|a\|} = \frac{1}{n} \to 0,$$

so $\|T_n - 0\| \to 0 \implies T_n \to 0$ in the norm topology (and thus in all the other topologies).

(b) Let $S_n(a_1, \dots) = (0, \dots, 0, a_{n+1}, a_{n+2}, \dots)$. Then if $x = \{x_n\} \in \ell^2$,

$$\|S_n x\|^2 = \sum_{i=n+1}^{\infty} |x_i|^2 \to 0,$$

so $S_n x \to 0\ \forall x \in \ell^2$, so $S_n \to 0$ as an operator from $\ell^2 \to \ell^1$ in the strong toplogy (and also the weak topology), but not the norm topology.

(c) Let $W_n : \ell^2 \to \ell^1$ be defined by $W_n(a_1, a_2, \dots) = (0, \dots, 0, a_1, a_2, \dots)$ with $n$ zeroes in the first position. Then if $x = (x_1, x_2, \dots) \in \ell^2, \ell \in \ell^{2^*}$, $\ell(W_n x) = \ell(0, \dots, 0, x_1, x_2, \dots)$. Recall that $\ell^2$ is a Hilbert space. We can show that $W_n \to 0$ in the weak operator topology.

**Theorem 7.5.** *Let $\mathcal{H}$ be a Hilbert space. Let $T_n \in \mathcal{L}(\mathcal{H}, \mathcal{H})\ \forall n$. Suppose $\forall x, y \in \mathcal{H}$, we have that $\langle x, T_n y\rangle$ converges as $n \to \infty$. Then $\exists T \in \mathcal{L}(\mathcal{H}, \mathcal{H})$ s.t.*

$$T_n \xrightarrow{w} T,$$

*where $\xrightarrow{w}$ denotes weak convergence.*

### 7.2 Adjoint Operators

Let $X, Y$ be Banach spaces and $T : X \to Y$ be linear and continuous. Define the *adjoint* operator $T' : Y^* \to X^*$ by $(T'\ell)x = \ell(Tx)$. Note that we want to think of $T'\ell \in X^*$.

**Example 7.6.** Let us revisit the shift operator. Let $X = Y = \ell_1$ and $T(a_1, a_2, \dots) = (0, a_1, a_2, \dots)$. Then $T' : \ell_1^* \to \ell_1^*$, or $T' : \ell_\infty \to \ell_\infty$.

**Claim 7.7.** *If $b \in \ell_\infty, b = (b_1, b_2, \dots)$, then $T'b = (b_2, b_3, \dots)$.*

**Theorem 7.8.** *Let $X, Y$ be Banach spaces. Then if $T : X \to Y$ is linear and bounded, the map $T \to T'$ which takes $T$ to its adjoint*

*(i) linear,*

*(ii) isometric, i.e., $\|T\| = \|T'\|$.*

Now suppose $\mathcal{H}$ is a Hilbert space. Let $T : \mathcal{H} \to \mathcal{H}$ be a bounded linear operator. The *Hilbert space adjoint* of $T$ is a bounded linear operator $T^*$ s.t. $\forall x, y \in \mathcal{H}$, $\langle x, Ty\rangle = \langle T^* x, y\rangle$.

Recall that if $x \in \mathcal{H}, \exists \ell \in \mathcal{H}^*$ st. $\forall y \in \mathcal{H}$, $\langle x, y\rangle = \ell(y)$. Let $C$ be the map defined by $x \mapsto \ell$, so $\|x\| = \|\ell\|$. If $x$ is replaced by $\alpha x$ for $\alpha \in \mathbb{C}$, then $\ell$ is replaced by $\overline{\alpha}\ell$. Thus, $C(\alpha x) = \overline{\alpha}\ell = \overline{\alpha}C(x)$, so $C$ is *conjugate linear.*

**Remark 7.9.** If $x, y \in \mathcal{H}$, $\langle T^* x, y\rangle = (C(T^* x))y$, while $\langle x, Ty\rangle = (Cx)(Ty) = (T'Cx)(y)\ \forall x, y \in \mathcal{H}$, so $CT^* = T'C$, so $T^* = C^{-1}T'C$. This is the relation between the Hilbert space adjoint and the regular adjoint operator.

If $\mathcal{H}$ is a Hilbert space, let $\mathcal{L}(\mathcal{H})$ be the bounded linear operators from $\mathcal{H}$ to $\mathcal{H}$.

**Theorem 7.10.** *We have the following:*

*(a) $T \to T^*$ is a conjugate linear isometric isomorphism of $\mathcal{L}$.*

*(b) For $S, T \in \mathcal{L}(\mathcal{H})$, define $ST = S \circ T$ to be the composition of $S$ and $T$. Then $(ST)^* = T^* S^*$.*

*(c) $(T^*)^* = T$.*

*(d) If the inverse map $T^{-1}$ exists and is a bounded linear operator, then $T^*$ also has a bounded inverse, and*

$$(T^*)^{-1} = (T^{-1})^*.$$

*(e) The map $T \to T^*$ is continuous in the weak operator topology and the norm operator topology, but not the strong operator topology.*

*(f) $\|T^* T\| = \|T\|^2$.*

The operator $T : \mathcal{H} \to \mathcal{H}$ is *self-adjoint* if $T^* = T$.

### 7.3 Complex Analysis

Let $f : \mathbb{C} \to \mathbb{C}$. Suppose $\Gamma \subset \mathbb{C}$ is a curve from point $a$ to point $b$. Suppose we mark of points $a = z_0, z_1, \ldots, z_n = b$ and define $\Delta z_i = z_i - z_{i-1}$. Define the integral

$$\int_\Gamma f(z)\, dz = \lim_{\substack{\Delta z_i \to 0 \\ n\to\infty}} \sum_{i=1}^n f(z_i)\Delta z_i.$$

Note that the contour has an *orientation.*

A set $D \subset \mathbb{R}^n$ is *connected* if any two points in $D$ can be connected by a broken line entirely inside of $D$. Suppose $D \subset \mathbb{C}$ is a connected open set. Then $f$ is *analytic* (*differentiable*)) in $D$ if $\forall z \in D$,

$$f'(z) := \lim_{h\to 0} \frac{f(z+h) - f(z)}{h}$$

exists, where $f'(z)$ is the *derivative* of $f(z)$.

**Lemma 7.11** (Cauchy's formula)**.** *Suppose $f : D \to \mathbb{C}$ is analytic. Let $\Gamma \subset D$ be a closed curve, oriented counterclockwise. Let $z_0 \in \Gamma$. Then*

$$f(z_0) = \frac{1}{2\pi i} \int_\Gamma \frac{f(z)}{z - z_0}\, dz.$$

Let $f : D \to \mathbb{C}$ be analytic, and let $z_0 \in D$. Suppose $f^{(n)}$ is the $n$th derivative of $f$ and let $a_n = f^{(n)}(z_0)/n!$. Then

$$\sum_{n=0}^\infty a_n(z - z_0)^n = f(z)$$

in the largest circular disk $C$ about $z_0$ in which $f$ can be extended as an analytic function. Conversely, every power series of the form

$$\sum_{n=0}^\infty a_n(z - z_0)^n$$

converges to an analytic function $f$ inside some circular disk $C$ about $z_0$, and diverges outside $C$. The function $f$ which is obtained cannot be extended as an analytic function to a larger concentric disk than $C$.

### 7.4 Banach-Valued Functions

Let $D \subset \mathbb{C}$ and $X$ be a complex Banach space. The function $f : D \to X$ is *analytic* if

$$f'(z) := \lim_{h\to\infty} \frac{f(z+h) - f(z)}{h}$$

exists $\forall z \in D$, where the limit is in the topology of $X$.

Let $D \subset \mathbb{C}$, $X$ be a Banach space, and $f : D \to X$. Then $f$ is *weakly analytic* if $\forall \ell \in X^*$,

$$\ell(f(z)) : D \to \mathbb{C}$$

is an analytic function.

**Lemma 7.12.** *Let $X$ be a Banach space. Then a sequence $\{x_n\} \subset X$ is Cauchy iff $\forall \ell \in X^*$ s.t. $\|\ell\| \leq 1$, $\{\ell(x_n)\}$ is Cauchy uniformly in $\ell \in X^*$.*

**Remark 7.13.** Uniformly in $\ell$ means that $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ s.t. $\forall m, n > N$, $|\ell(x_n) - \ell(x_m)| < \varepsilon$, *independent of* $\ell$.

**Theorem 7.14.** *If $f : D \to X$ is weakly analytic, then $f$ is strongly analytic.*

### 7.5 The Spectrum of a Linear Operator

Let $X$ be a Banach space. The operator $I : X \to X$ denotes the *identity*: $Ix = x\ \forall x \in X$. Define $\mathcal{L}(X)$ to be all bounded linear operators from $X \to X$.

Let $T : X \to X$ be bounded and linear and $\lambda \in \mathbb{C}$. Then $\lambda$ is in the *resolvent set* of $T$ if the operator $\lambda I - T$ has an inverse $(\lambda I - T)^{-1}$ which is bounded and linear. Let $\rho(T)$ denote the resolvent set of $T$. Let $R_\lambda(T) := (\lambda I - T)^{-1}$. We also write $(\lambda I - T)^{-1} = (\lambda - T)^{-1} = \frac{1}{\lambda - T}$.

The product $AB$ of two operators $A, B$ denotes their composition. Two operators $A, B$ *commute* if $AB = BA$.

**Lemma 7.15.** *If the operator $A \in \mathcal{L}(X)$ has norm smaller than 1, then $1/(1-A)$ exists and*

$$\frac{1}{1-A} = \sum_{k=0}^{\infty} A^k.$$

**Theorem 7.16.** *Suppose $T$ is a bounded linear operator. We have the following:*

*(a) $\rho(T) \subset \mathbb{C}$ is open.*

*(b) For $\lambda \in \rho(T)$, $R_\lambda(T) = (\lambda - T)^{-1}$ is an analytic function of $\lambda$, i.e., $R_\lambda(T) : \mathbb{C} \to \mathcal{L}(X)$ is an analytic function of $\lambda$.*

*(c) Further, for $\lambda, \mu \in \mathbb{C}$,*

$$R_\lambda(T) - R_\mu(T) = (\mu - \lambda) R_\mu(T) R_\lambda(T),$$

*and $R_\lambda(T), R_\mu(T)$ commute.*

**Remark 7.17.** If $X = \mathbb{R}^n$ and $T$ is a matrix, then

$$(\lambda I - T)^{-1} - (\mu I - T)^{-1} = (\mu - \lambda)(\mu I - T)^{-1}(\lambda I - T)^{-1}.$$

The *spectrum* $\sigma(T)$ of the operator $T$ is the set of $\lambda \in \mathbb{C}$ s.t. $R_\lambda(T)$ does not exist as a bounded operator, *i.e.*, $\sigma(T) = \mathbb{C} \setminus \rho(T)$.

**Remark 7.18.** By part (a) of Theorem 7.16, the spectrum of $T$ is a closed set.

Let $X$ be a Banach space. A function $f(\lambda) : \mathbb{C} \to X$ is called *entire* if it is analytic $\forall \lambda \in \mathbb{C}$.

**Theorem 7.19** (Liouville's theorem for Banach spaces)**.** *Let $X$ be a Banach space. Let $f(\lambda) : \mathbb{C} \to X$ be entire. Then if $f(\lambda)$ is bounded, i.e., if $\|f(\lambda)\| \le C \ \forall \lambda \in \mathbb{C}$, then $f(\lambda)$ is a constant, i.e., $\exists T \in X$ s.t. $f(\lambda) = T \ \forall \lambda \in \mathbb{C}$.*

**Corollary 7.20.** *Let $X$ be a Banach space, $T \in \mathcal{L}(X)$. We have the following:*

*(a) The spectrum $\sigma(T)$ is non-empty.*

*(b) Further, $\sigma(T) \subset D$, a closed disk of radius $\|T\|$.*

Could it be possible that $\sigma(T) \subset D'$ a closed disk of radius $\|T\| - \varepsilon$ for some $\varepsilon > 0$? The answer is no.

Define the *spectral radius* of $T$ be $r(T) := \sup_{\lambda \in \sigma(T)} |\lambda|$ to be the largest distance of any point in $\sigma(T)$ from 0. Given an analytic function $f : \mathbb{C} \to \mathbb{C}$ and its power series about $a \in \mathbb{C}$:

$$f(z) = \sum_{k=0}^{\infty} c_n (z-a)^n, \tag{8}$$

its *radius of convergence* $\rho$ is defined to be the smallest $r$ s.t. the series in Equation (8) converges for $|z - a| < r$.

**Lemma 7.21.** *Given an analytic function $f$ with its Taylor series in Equation* (8)*, the radius of convergence $R$ satisfies*

$$R = \limsup_{n \to \infty} |a_n|^{1/n}.$$

**Theorem 7.22.** *We have the following.*

1. *Given a Banach space $X$ and bounded operator $T$ on $X$, the spectral radius $r(T)$ is given by*

$$r(T) := \lim_{n \to \infty} \|T^n\|^{1/n} := |||T|||.$$

2. *If $X$ is a Hilbert space and $T$ is self-adjoint, then*

$$r(T) := \|T\|.$$

**Theorem 7.23.** *Let $\mathcal{H}$ be a Hilbert space and $T$ be a bounded operator on $\mathcal{H}$. Then $\sigma(T^*) = \overline{\sigma(T)} = \{\overline{\lambda} \mid \lambda \in \sigma(T)\}$ and $R_{\overline{\lambda}}(T^*) = R_\lambda(T)^*$ if $\lambda \in \rho(T)$.*

### 7.6 Eigenvalues and the Point Spectrum

Let $T$ be a bounded operator on a Banach space $X$. Suppose $\lambda \in C$, and for some $x \in X$, $Tx = \lambda x$. Then $\lambda$ is an *eigenvalue* of $T$. If $\lambda \in \sigma(T)$ is an eigenvalue of $T$, then $\lambda$ is in the *point spectrum* of $T$.

**Remark 7.24.** In finite dimensions, the point spectrum is always the entire spectrum. In general, the point spectrum is not the full spectrum. In infinite dimensions, there are additional points in the spectrum, *e.g.*, the "continuous spectrum" and the "residual spectrum." We will not focus so much on the residual spectrum.

**Theorem 7.25.** *Let $\mathcal{H}$ be a Hilbert space. If $T : \mathcal{H} \to \mathcal{H}$ is self-adjoint, then*

*(a)* $\sigma(T) \subset \mathbb{R} \subset \mathbb{C}$.

*(b) If $\lambda_1, \lambda_2$ are eigenvalues and $Tx_1 = \lambda_1 x_1, Tx_2 = \lambda_2 x_2$, then $x_1 \perp x_2$.*

### 7.7 Polar Decomposition

Let $\mathcal{H}$ be a Hilbert space. An operator $A \in \mathcal{L}(\mathcal{H})$ is *positive* if $\langle Ax, x\rangle = 0\ \forall x \in \mathcal{H}$.

**Proposition 7.26.** *If $A$ is positive, then $A$ is self-adjoint.*

**Lemma 7.27** (Square root lemma)**.** *Let $\mathcal{H}$ be a Hilbert space and $A \in \mathcal{L}(\mathcal{H})$ s.t. $A \geq 0$. Then $\exists! B \in \mathcal{L}(\mathcal{H})$ with $B \geq 0$ s.t. $B^2 = A$. Further, the operator $B$ commutes with every bounded operator that commutes with $A$.*

Let $A \in \mathcal{L}(\mathcal{H})$. Then define $|A| = \sqrt{A^*A}$.

If $T \in \mathcal{L}(\mathcal{H})$, the *kernel* of $T$ is defined by $\ker T := \{x \mid Tx = 0\}$. Let $U \in \mathcal{L}(\mathcal{H})$. Then $U$ is an *isometry* if $\|Ux\| = \|x\|\ \forall x \in \mathcal{H}$. $U$ is a *partial isometry* if $\|Ux\| = \|x\|\ \forall x \in (\ker U)^\perp$.

**Remark 7.28.** If $U : \mathcal{H} \to \mathcal{H}$ is a partial isometry, note that $\mathcal{H} = \ker U \oplus (\ker U)^\perp$ uniquely and $\mathcal{H} = \operatorname{ran} U \oplus (\operatorname{ran} U)^\perp$ uniquely.

**Remark 7.29.** Note that $\operatorname{ran} U|_{(\ker U)^\perp} = \operatorname{ran} U$, so we can say $U : (\ker U)^\perp \to \operatorname{ran} U$.

Recall that for a $z \in \mathbb{C}$, we can write $z = re^{i\theta} = |z|e^{i\theta}$. Note that we can interpret $z$ as as a multiplication operator $M_z$, that is if $h \in \mathcal{H}$, we define $M_z h = zh$. This also generalizes to more complicated operators:

**Theorem 7.30** (Polar decomposition theorem)**.** *Let $A \in \mathcal{L}(\mathcal{H})$. Then there exists partial isometry $U$ s.t. $A = U|A|$. $U$ is unique if it is chosen so that* $\ker U = \ker A$. *Also,* $\operatorname{ran} U = \overline{\operatorname{ran} A}$.

# 8 The Spectral Theorem

### 8.1 Functional Calculus

Let $\mathcal{H}$ be a complex Hilbert space. Let $A \in \mathcal{L}(\mathcal{H})$. Recall that $A^2 = A \circ A$, $A^3 = A \circ A \circ A$, etc. If $P(x) = \sum_N a_i x^i$ is a complex polynomial, $P(A) := \sum_N a_i A^i$ is an operator. Let $f(x)$ be a continuous function. Is $f(A)$ an operator as well?

The answer is yes! Use an approximation of $f(x)$ by $P(x)$. Note that $P(x) \approx f(x)$ is only needed on $\sigma(A)$.

Define $\phi : C(\sigma(A)) \to \mathcal{L}(\mathcal{H})$. We choose $\phi$ with the following properties:

(a) $\phi$ is a $*$-homomorphism, i.e.

  (i) $\phi(f + g) = \phi(f) + \phi(g)$,

  (ii) $\phi(fg) = \phi(f)\phi(g)$,

  (iii) $\phi(\lambda f) = \lambda\phi(f)$,

  (iv) $\phi(1) = I$,

  (v) $\phi(\overline{f}) = \phi(f)^*$.

(b) $\phi$ is continuous from $C(\sigma(A))$ (in sup norm) to $\mathcal{L}(\mathcal{H})$ (in operator norm) for fixed $A$.

(c) Let $f(x) = x$ on $\mathbb{R}$. Then $\phi(f) = A$.]

Moreover, $\phi$ has the additional properties:

(d) If $A\psi = \lambda\psi$, then $\phi(f)\psi = f(\lambda)\psi$.

(e) $\sigma[\phi(f)] = \{f(\lambda \mid \lambda \in \sigma(A)\}$ [spectral mapping theorem].

(f) If $f \geq 0$, then $\phi(f) \geq 0$.

(g) $\|\phi(f)\| = \|f\|_\infty$ [this strengthens (b)].

### 8.2 The Spectral Mapping Theorem

**Lemma 8.1** (Spectral mapping theorem for polynomials)**.** *For a polynomial $P(x)$ and operator $A \in \mathcal{L}(\mathcal{H})$, we have*

$$\sigma(P(A)) = P(\sigma(A)) := \{P(\lambda) \mid \lambda \in \sigma(A)\}.$$

**Lemma 8.2.** *For a polynomial $P$,*

$$\|P(x)\|_{C(\sigma(A))} = \|P(A)\|_{\mathcal{L}(\mathcal{H})}.$$

### 8.3 The Borel Functional Calculus

Again, consider a fixed bounded operator $A$. If $f : \mathbb{R} \to \mathbb{C}$ is Borel, can we define $f(A)$? This means we want the correspondence $\hat{\phi} : f \to f(A)$ which extends $\phi$ (which was defined only for continuous functions).

**Theorem 8.3.** *A unique correspondence $\hat{\phi}$ exists between Borel functions $f(x)$ and operators $f(A)$ and extends the correspondence $\phi$ above. Further, if we require that $\hat{\phi}$ satisfies the same three conditions (a)-(c) as $\phi$, and that if $f_n(x) \to f(x)$ $\forall x$, and $\|f_n\|_\infty \leq C$ $\forall n$, then $f_n(A) \to f(A)$ in the strong operator topology.*

### 8.4 Spectral Measures

Let $\mathcal{H}$ be a complex Hilbert space and let $A \in \mathcal{L}(\mathcal{H})$ be a fixed self-adjoint operator. Then $\sigma(A) \subset \mathbb{R}$. Fix $\psi \in \mathcal{H}$, and ffor each continuous function $f$ on $\sigma(A)$, define $\ell(f) = \langle \psi, f(A)\psi \rangle$, which is a number. Here, $A, \psi$ are *fixed* as $f$ varies.

It is easy to show that if $f \geq 0$, then $f(x) = g^2(x)$ for some $g \geq 0$, so $\ell(f) = \langle \psi, g^2(A)\psi \rangle = \langle g(A)\psi, g(A)\psi \rangle \geq 0$. Hence, $\ell$ is a positive bounded linear functional on functions $f$ on $\sigma(A)$. Then by the Riesz-Markov theorem, there exists a Baire measure $\mu_\psi$ on $\sigma(A)$ s.t.

$$\ell(f) = \int_{\sigma(A)} f(x)\, d\mu_\psi.$$

$\mu_\psi$ is the *spectral measure* associated with $\psi$.

A vector $\psi \in \mathcal{H}$ is a *cyclic vector* for $A$ if finite linear combinations of the elements $\{A^n\psi\}$ span $\mathcal{H}$, *i.e.*, finite combinations of vectors $A^n\psi$ are dense in $\mathcal{H}$.

**Example 8.4.** Let $\mathcal{H} = L^2[-1, 1]$ and $A = M_x$, *i.e.*, $Af(x) = M_x f(x) = xf(x)$. Let $\psi = 1$. Then $\{A^n 1\} = \{x^n\}$. The finite linear combinations of $\{A^n 1\}$ are thus polynomials, which are dense in $\mathcal{H}$. Thus, the vector $\psi = 1$ is a cyclic vector for the operator $A = M_x$.

**Claim 8.5.** *If $A$ has a cyclic vector $\psi$, then it is easy to describe $A$ as an operator.*

Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces. Suppose $A_1 \in \mathcal{L}(\mathcal{H}_1), A_2 \in \mathcal{L}(\mathcal{H}_2)$. Suppose that there exists unitary operator $U : \mathcal{H}_1 \to \mathcal{H}_2$ and suppose $\forall x \in \mathcal{H}_2$, $A_2 x = UA_1U^{-1}x$, then $A_1$ and $A_2$ are *unitarily equivalent.*

Now let $\mathcal{H}$ be a Hilbert space and $A$ be a self-adjoint operator on $\mathcal{H}$. Let $\psi$ be a cyclic vector for $A$. Suppose $\mu_\psi$ is the corresponding measure on $\sigma(A)$. Consider $L^2(\sigma(A), \mu_\psi)$ as a Hilbert space. Let $A_1$ be a self-adjoint bounded operator on $L^2(\sigma(A), \mu_\psi)$ be defined by $(A_1 f)(x) = xf(x)$.

**Lemma 8.6.** *$A$ and $A_1$ are unitarily equivalent. That is, there exists unitary operator $U : \mathcal{H} \to L^2(\sigma(A))$ s.t. $UAU^{-1} = A_1$.*

**Lemma 8.7.** *Let $\mathcal{H}$ be a separable Hilbert space and $A \in \mathcal{L}(\mathcal{H})$ be self-adjoint and fixed. There exists orthogonal subspaces $\{\mathcal{H}_i\}_{i=1}^N$ of $\mathcal{H}$ ($N$ may be $\infty$) s.t.*

*(i)* $\mathcal{H} = \bigoplus \mathcal{H}_i$.

*(ii)* *If $\psi_i \in \mathcal{H}_i$, then $A\psi_i \in \mathcal{H}_i$.*

*(iii)* *$\forall i$, consider $A|_{\mathcal{H}_i}$. This operator has a cyclic vector $\psi_i \in \mathcal{H}_i$, i.e., finite linear combinations of $\{A^n\psi_i\}_{n=0}^\infty$ are dense in $\mathcal{H}_i$.*

Combining the above information:

1. If $A$ is self-adjoint on $\mathcal{H}$ with a cyclic vector $\psi$, then there is a measure $\mu$ on $\mathbb{R}$ (actually restricted to $\sigma(A) \subset \mathbb{R}$) s.t. $\mathcal{H} \leftrightarrow L^2(\mathbb{R}, \mu)$, *i.e.*, $\exists! U : \mathcal{H} \to L^2(\mathbb{R}, \mu)$.
2. Under the unitary operator $U$, the operator $A$ on $\mathcal{H}$ corresponds to $M_x$ on $L^2(\mathbb{R}, \mu)$, *i.e.*, $UAU^{-1} = M_x$.
3. If $A$ is a general self-adjoint operator on $\mathcal{H}$, then we can decompose $\mathcal{H}$ orthogonally: $\mathcal{H} = \bigoplus \mathcal{H}_n$ s.t. $A$ maps each $\mathcal{H}_n \subset \mathcal{H}$ into itself and $A|_{\mathcal{H}_n}$ has a cyclic vector.

Conclusion: On $\mathcal{H}_n$, $A$ has a cyclic vector $\psi_n$ and so there is a corresponding measure $\mu_n$ on $\mathbb{R}$ with $\mathcal{H}_n \leftrightarrow L^2(\mathbb{R}, \mu_n)$, *i.e.*, each piece of $A$ on each $\mathcal{H}_n$ is like the operation of mulitplying a function in $L^2(\mathbb{R}, \mu_n)$ by $x$, with a corresponding measure $\mu_n$. Combining all of this gives $A \leftrightarrow M_x$ acting on $\bigoplus L^2(\mathbb{R}, \mu_n)$.

But what is $\bigoplus L^2(\mathbb{R}, \mu_n)$? It is easy to show that $\bigoplus L^2(\mathbb{R}, \mu_n) = L^2(\cup(\mathbb{R}, \mu_n))$. So what kind of measure space is $\mathcal{M} := \cup(\mathbb{R}, \mu_n)$?

$\mathcal{M}$ is a *simple union* of separate copies of $\mathbb{R}$ with different measures $\mu_n$. Gluing them all together yields a single measure space. Thus, if $f(x)$ is a function on this measure space, it is defined (maybe differently) on each copy of $\mathbb{R}$. Thus, every self-adjoint $A$ is equivalent to operator $M_x$ on $L^2(\mathcal{M})$.

**Remark 8.8.** The equivalence still holds even if we replace $\mu_n$ by $c\mu_n$, where $c$ is dependent on $n$.

The measures $\mu_n$ are *spectral measures* of $A$.

**Example 8.9.** Let $\mathcal{H}$ be a finite-dimensional space, $A$ be self-adjoint on $\mathcal{H}$. Then $\mathcal{H} = \mathbb{C}^n$, and so $A$ can be represented as a matrix. Let $\{\lambda_i\}, \{\psi_i\}$ be eigenvalues and eigenvectors of $A$. Then $\sigma(A) = \{\lambda_i\}$ is a finite set of points. Assume that all $\lambda_i$ are unique. The spectral measure $\mu$ must be concentrated on the eigenvalues. Suppose $\mu$ is the counting measure on $\{\lambda_i\}$. Note that then we have

$$\int_{\mathbb{R}} f(x)\, d\mu = \sum f(\lambda_i).$$

Therefore, $\forall f(x) \in L^2(\mathbb{R}, \mu)$ is uniquely determined by its values on $\lambda_i$:

$$f \leftrightarrow (f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)),$$

*i.e.*, $L^2(\mathbb{R}, d\mu) \sim \mathbb{C}^n = \mathcal{H}$. Further, $A$ is equivalent to multiplication $M_x$ on $L^2(\mathbb{R}, d\mu)$.

### 8.5 More on the Spectral Theorem

**Lemma 8.10.** *On the space $L^2(\mathcal{M})$, the operator $M_{a(x)}$ with $a(x)$ essentially bounded is a bounded linear operator. If $a(x)$ is essentially unbounded, then $M_{a(x)}$ is not bounded as an operator.*

**Proposition 8.11.** *Let $A \approx M_x$ on $\bigoplus L^2(\mathbb{R}, \mu_n)$. Then $\sigma(A) = \overline{\cup \operatorname{supp} \mu_n}$.*

**Theorem 8.12.** *If $A$ is self-adjoint on $\mathcal{H}$, then $\mathcal{H} = \mathcal{H}_{pp} \oplus \mathcal{H}_{ac} \oplus \mathcal{H}_{sc}$, where if $\psi \in \mathcal{H}_{pp}$, then $\mu_\psi$ is a pure point measure on $\mathbb{R}$, $\psi \in \mathcal{H}_{ac}$, then $\mu_\psi$ is absolutely continuous on $\mathbb{R}$, and if $\psi \in \mathcal{H}_{sc}$, then $\mu_\psi$ is singular continuous on $\mathbb{R}$.*

### 8.6 Multiplicity: Free Operators

Recall if $A$ is self-adjoint on $\mathcal{H}$, then $A \sim M_x$ on $\mathcal{M}$, which is the union of copies $\mathbb{R}_n$ of $\mathbb{R}$ with some measure $\mu_n$. If $\mathcal{H}$ is finite-dimensional, then $\mu$ is always a point measure on $\mathcal{M}$.

f $A \sim M_x$ on some measure space $\mathcal{M}$ that is only one copy of $\mathbb{R}$ with some measure $\mu$, then $A$ is *multiplicity-free*.

**Theorem 8.13.** *The following are equivalent:*

*(a) $A$ is multiplicity-free,*

*(b) $A$ has a cyclic vector.*

**Proposition 8.14.** *Let $\mu, \nu$ be Borel measures on $\mathbb{R}$ with bounded support. Let*

$$\begin{aligned} A_\mu &= M_x \quad \text{on } L^2(\mathbb{R}, \mu), \\ A_\nu &= M_x \quad \text{on } L^2(\mathbb{R}, \nu). \end{aligned}$$

*Then $A_\mu$ and $A_\nu$ are unitarily equivalent iff $\mu, \nu$ are equivalent measures.*

Let $\Omega$ be a Borel set in $\mathbb{R}$. Let $\chi_\Omega$ be the indicator function for $\Omega$. If $A$ is self-adjoint on $\mathcal{H}$, then define (using functional calculus for operators) $P(\Omega) = \chi_\Omega(A)$, a *spectral projection* of $A$. By properties of functions of operators, we have

$$P(\Omega)^2 = \chi_\Omega^2(A) = \chi_\Omega(A).$$

Also,

$$P(\Omega)^* = \overline{\chi}_\Omega(A) = \chi_\Omega(A) = P(\Omega),$$

so $P(\Omega)$ is an orthogonal projection.

**Proposition 8.15.** *We have the following:*

- *(a)* $P_\Omega$ *is an orthogonal projection.*
- *(b)* $P(\phi) = 0$*;* $P(-a, a) = I$ *for a sufficiently large.*
- *(c) If* $\Omega = \cup_{m=1}^\infty \Omega_m$ *with* $\Omega_m$ *disjoint, then*

$$P_\Omega = \lim_{M\to\infty} \sum_{m=1}^{M} P(\Omega_m)$$

*in the strong operator topology.*

- *(d)* $P(\Omega_1)P(\Omega_2) = P(\Omega_1 \cap \Omega_2)$.

If $\forall \Omega \subset \mathbb{R}$, $P(\Omega)$ is an orthogonal projection on $\mathcal{H}$ and properties (a)-(d) are satisfied, then $P$ is a *projection-valued measure*.

Suppose $\theta \in \mathcal{H}$ is fixed. Define a measure $\mu$ on $\mathbb{R}$ by

$$\mu(\Omega) = \langle \theta, P(\Omega)\theta \rangle.$$

We can define an integral w.r.t. $P(\Omega)$:

**Theorem 8.16.** *If* $P_\Omega$ *is a projection-valued measure and* $f(x)$ *is a bounded Borel function on* $\mathbb{R}$*, there exists a unique operator* $B$ *s.t.* $\forall \theta \in \mathcal{H}$,

$$\langle \theta, B\theta \rangle = \int f(x)\, d\langle \theta, P(x)\theta \rangle.$$

*Define*

$$B = \int f(x)\, dP(x).$$

We can show if $\{P(\Omega)\}$ are projection-valued measures associated with $A$, then

$$f(A) = \int_{\mathbb{R}} f(x)\, dP(x)$$

s.t.

$$A = \int_{\mathbb{R}} x\, dP(x).$$

**Theorem 8.17** (Spectral theorem)**.** *There is a one-to-one correspondence between self-adjoint operators* $A$ *and projection-valued measure* $P(\Omega)$ *defined by* $\theta_1 : A \to \{P(\Omega)\}$*, a projection-valued measure defined by* $A$ *and* $\theta_2 : \{P(\Omega)\} \to A = \int_{\mathbb{R}} x\, dP(x)$.