Functional Analysis

Jonathan Ma
johnma@udel.edu

January 24, 2026

Contents

Preface

Adapted from MIT 18.102, Introduction to Functional Analysis, Spring 2021: found here. Fair warning, the content is undergrad level, and does not assume measure theory knowledge, so there is a large portion of the course dedicated towards developing the Lebesgue integral.

1 Lecture 1. Basic Banach Space Theory

The prereqs for this course are linear algebra and calculus (real analysis). These courses work in finite dimensional spaces. Typically, we are concerned with solving finite systems of equations, or in general solving problems with finitely many independent variables. But ODEs, PDEs typically don't have sets of finite independent variables. Suppose the independent variables we work with are members of some vector space. In calculus, we have functions on $\mathbb{R}, \mathbb{R}^2, \mathbb{R}^3$, etc. Functional analysis was built to study cases when the vector spaces we work with are not necessarily finite dimensional.

1.1 Normed Spaces

Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$. (We will denote these fields with $\mathbb{K}$.) In analysis and linear algebra, we were introduced to (perhaps) two notions of the size of a space. One notion was cardinality. The other notion was dimension.

Definition 1.1. Let V be a vector space. If every linearly independent set is finite, then we say V is called finite dimensional. Some books also define V as finite dimensional if it has a finite basis. We say V is infinite dimensional if V is not finite dimensional.

Example 1.2. The vector spaces $\mathbb{R}, \mathbb{R}^n, \mathbb{C}^n$, are finite dimensional. The vector space $C[0,1]$ is infinite dimensional. Recall that $C[0,1]$ is the space of continuous functions on $[0,1]$. Consider that $E = \{f_n(x) = x^n \mid n \in \mathbb{Z}^+ \cup \{0\}\}$ is a linearly independent set with infinite cardinality.

In analysis, we proved the Heine-Borel theorem, that is, closed and bounded subsets of $\mathbb{R}^n$ are compact. (Every bounded subsequence has a convergent subsequence.) Heine-Borel is then used to prove that every continuous function on a closed and bounded set has a min and a max in that set. This statement is generally not true in the infinite dimensional setting. Doing analysis becomes harder in the infinite dimensional setting, because these nice properties about finite dimensional spaces do not hold true in the infinite dimensional setting. To do analysis, we now need some notion of how close things in a space are.

Definition 1.3 (Norm, Normed Vector Space)**.** A norm on a vector space V is a function $\|\cdot\| : V \to [0,\infty)$ satisfying:

1. $\|v\| = 0 \iff v = 0, \quad \forall v \in V$ (definiteness),
2. $\|\lambda v\| = |\lambda|\|v\|, \quad \forall v \in V,\ \forall \lambda \in \mathbb{K}$ (homogeneity),
3. $\|v + w\| \leq \|v\| + \|w\|, \quad \forall v, w \in V$ (triangle inequality).

We call V a normed vector space if it is equipped with a norm.

Definition 1.4 (Seminorm)**.** A semi-norm is a function $\|\cdot\| : V \to [0,\infty)$ satisfying homogeneity and the triangle inequality, but not necessarily definiteness.

Definition 1.5 (Metric, Metric Space)**.** Let X be a set. Recall that a function $d : X \times X \to [0,\infty)$ is a metric if

1. $d(x,y) = 0 \iff x = y, \quad \forall x, y \in X$,
2. $d(x,y) = d(y,x), \quad \forall x, y \in X$,
3. $d(x,y) \leq d(x,z) + d(z,y), \quad \forall x, y, z \in X$.

We call (X,d) a metric space in this case.

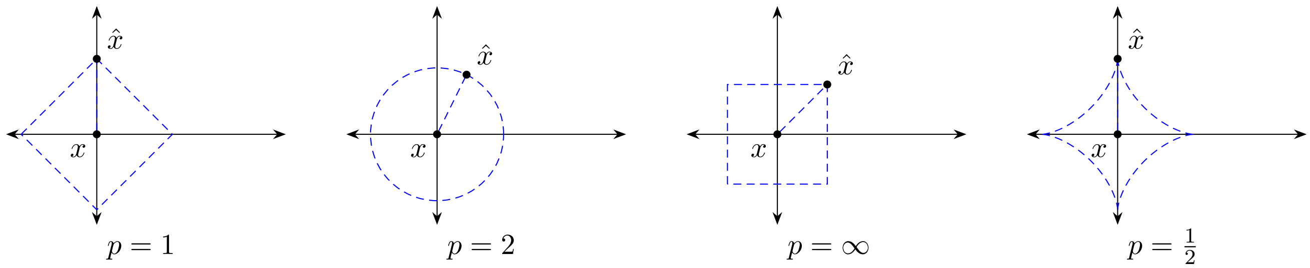


Figure 1: Unit balls in $\mathbb{R}^2$ under various ℓ^p norms.

Theorem 1.6. *If V is a normed vector space with norm $\|\cdot\|$, then*

$$d(v, w) := \|v - w\|, \quad \forall v, w \in V,$$

defines a metric on V.

Proof. The first property of norms implies the first property of metric spaces. Let $v, w \in V$. Then

$$\|v - w\| = \|-1(w - v)\| = |-1|\|w - v\| = \|w - v\|,$$

so the second property of metric spaces holds. The third property of metric spaces also falls immediately from the third property of norms. $\square$

Let's look at some norms.

Example 1.7. Consider $\mathbb{R}^n$ with the Euclidean norm: for $x \in \mathbb{R}^n$,

$$\|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}.$$

We also have the infinity norm:

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

More generally,

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, \quad 1 \leq p < \infty,$$

is called the ℓ^p norm ("little ell-p norm"). One can prove that for a fixed vector $x \in \mathbb{R}^n$, as we take p to infinity, the ℓ^p norm "converges" to the infinity norm.

Figure 1 demonstrates how certain unit balls in $\mathbb{R}^2$ are graphed under certain ℓ^p norms.

Example 1.8. Let (X, d) be a metric space. Define $C_\infty(X) = \{f : X \to \mathbb{C} \mid f \text{ is continuous and bounded}\}$. For example, if we take $X = \mathbb{R}^n$ then $C_\infty[0, 1] = C[0, 1]$. (Since $[0, 1]$ is compact.) Then $C_\infty(X)$ is a vector space. We claim

$$\|u\|_\infty = \sup_{x \in X} \{|u(x)|\}, \quad \forall u \in X,$$

is a norm on $C_\infty(X)$. This is easy enough to prove. Note that $u_n \to u$ in $C_\infty(X)$ iff $\|u_n - u\|_\infty \to 0$ as $n \to \infty$. That is, $\forall \varepsilon > 0, \exists N \in \mathbb{Z}^+$ such that

$$\forall n \in \mathbb{Z}^+, n \geq N \implies (\forall x \in X, \; |u_n(x) - u(x)| < \varepsilon).$$

But this holds iff $u_n \to u$ uniformly on X.

Example 1.9. We write $\{a_n\}$ to mean a sequence in $\mathbb{R}^n$. As an abuse of notation, we will also write $\{a_n\}_{n=1}^{\infty}$ to mean the same. Consider the ℓ^p space:

$$\ell^p = \{\{a_n\} \mid \|\{a_n\}\|_p < \infty\},$$

where

$$\|\{a_n\}\|_p = \left(\sum_{n=1}^{\infty} |a_n|^p\right)^{1/p}.$$

Also, we write

$$\|\{a_n\}\|_\infty = \sup_{1 \leq n < \infty} \{|a_n|\}.$$

As an example, $\left\{\frac{1}{n}\right\}_{n=1}^{\infty} \in \ell^p, \forall p > 1$. But not for $p = \infty$. None of these facts about ℓ^p spaces (or the fact that ℓ^p spaces are vector spaces) are trivial to prove, and are course exercises.

The central objects of interest in functional analysis are the analogs of $\mathbb{R}^n, \mathbb{C}^n$. The nice thing about $\mathbb{R}^n, \mathbb{C}^n$ is that the metrics on these sets are complete. (Cauchy sequences always converge.)

Definition 1.10 (Banach Space)**.** A normed space V is a Banach space if it is complete with respect to the metric induced by its norm. Named for Stefan Banach (1892 – 1945).

Example 1.11. The sets $\mathbb{R}^n, \mathbb{C}^n$ are Banach spaces with respect to any $\|\cdot\|_p$ norms.

Let us show that $C_\infty(X)$ is a Banach space.

Theorem 1.12. *If X is a metric space, then $C_\infty(X)$ is a Banach space.*

Proof. We want to show that every Cauchy sequence $\{u_n\}$ in $C_\infty(X)$ has a limit in $C_\infty(X)$. Let $\{u_n\}$ be a Cauchy sequence in $C_\infty(X)$. Since $\{u_n\}$ is Cauchy, we claim it is bounded. Consider that there exists $N_0 \in \mathbb{Z}^+$ such that

$$\forall n, m \in \mathbb{Z}^+, \|u_n - u_m\|_\infty < 1.$$

Then for all $n \geq N_0$,

$$\begin{aligned} \|u_n\|_\infty &\leq \|u_n - u_{N_0}\|_\infty + \|u_{N_0}\|_\infty \\ &< 1 + \|u_{N_0}\|_\infty. \end{aligned}$$

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

$$\begin{aligned} \|u_n\|_\infty &\leq \|u_1\|_\infty + \|u_2\|_\infty + \ldots + \|u_{N_0}\|_\infty + 1 \\ &:= B. \end{aligned}$$

Now, let $\varepsilon > 0$ be given, and let $N \in \mathbb{Z}^+$ be such that

$$\forall n, m \in \mathbb{Z}^+, n, m \geq N \implies \|u_n - u_m\|_\infty < \varepsilon.$$

Fix an arbitrary $x \in X$. Then since

$$|u_n(x) - u_m(x)| \leq \|u_n - u_m\|_\infty,$$

we have that $\{u_n(x)\}_{n=1}^{\infty}$ is a Cauchy sequence. Since $\{u_n(x)\}$ is a sequence in $\mathbb{C}$, and it is Cauchy, we know it converges. Define $u : X \to \mathbb{C}$ such that

$$u(x) = \lim_{n\to\infty} u_n(x).$$

So far, we've shown that $u_n \to u$ pointwise. Consider that

$$\begin{aligned} |u(x)| &= \lim_{n\to\infty} |u_n(x)| \le B \\ &\implies \sup_{x\in X} |u(x)| \le B. \end{aligned}$$

Now we want to show $\|u - u_n\|_\infty \to 0$ as $n \to \infty$. Let $N' \in \mathbb{Z}^+$ be such that

$$\forall n, m \in \mathbb{Z}^+, n, m \ge N' \implies \|u_n - u_m\|_\infty < \frac{\varepsilon}{2}.$$

Fix $x \in X$ again, then suppose $s, t \in \mathbb{Z}^+, s, t \ge N'$. Then,

$$\begin{aligned} |u_s(x) - u_t(x)| &\le \|u_s - u_t\|_\infty < \frac{\varepsilon}{2} \\ \implies \forall n \ge N', |u_n(x) - u(x)| &\le \frac{\varepsilon}{2} \qquad \text{(as } n \to \infty.) \\ \implies \|u_n - u\|_\infty &\le \frac{\varepsilon}{2} < \varepsilon. \end{aligned}$$

Thus, $\|u_n - u\|_\infty \to 0$. We lastly need to show that $u \in C_\infty(X)$. We've already shown u is bounded. Since $u_n \to u$ uniformly, u must be continuous. Thus $u \in C_\infty(X)$, and thus, $C_\infty(X)$ is complete, thus, it is a Banach space. □

Exercise 1.1. Show that ℓ^p is a Banach space for all $1 \le p < \infty$.

Exercise 1.2. Show that $c_0 = \{\alpha \in \ell^\infty \mid \lim_{n\to\infty} \alpha_n = 0\}$ is a Banach space with the ℓ^∞ norm.

2 Lecture 2. Bounded Linear Operators

Let V be a normed vector space.

Definition 2.1. Let $\{v_n\} \subset V$. The series $\sum_{n=1}^{\infty} v_n$ is summable if $\{\sum_{n=1}^{m} v_n\}_{m=1}^{\infty}$ converges. We say $\sum_{n=1}^{\infty} v_n$ is absolutely summable if $\sum_{n=1}^{\infty} \|v_n\|$ converges. We also say $\sum_{n=1}^{\infty} v_n$ is convergent, or absolutely convergent, from earlier analysis courses.

Theorem 2.2. *If $\sum_{n=1}^{\infty} v_n$ is absolutely summable, then $\{\sum_{n=1}^{m} v_n\}_{m=1}^{\infty}$ is Cauchy in V.*

Proof. Proof is the same as in the $V = \mathbb{R}$ case. □

Note that this condition is weaker than what we proved in real analysis, since V is not necessarily complete. Now we will discuss a nice theorem which we may use to prove certain spaces are Banach.

Theorem 2.3. *A normed vector space V is Banach iff all absolutely summable series in the space are summable.*

Proof. Assume that V is Banach. Let $\{v_n\} \subset V$, and suppose $\sum_{n=1}^{\infty} v_n$ is absolutely summable. Then $\{\sum_{n=1}^{m} v_n\}_{m=1}^{\infty}$ is Cauchy, hence, convergent. Thus, $\sum_{n=1}^{\infty} v_n$ is summable. For the converse, assume that each absolutely summable series in V is summable. Let $\{v_n\} \subset V$ be Cauchy. We will show that $\{v_n\}$ has a convergent subsequence. (A Cauchy sequence is convergent iff it has a convergent subsequence.) Since $\{v_n\}$ is Cauchy, for all $k \in \mathbb{Z}^+$, there exists an $N_k \in \mathbb{Z}^+$ such that

$$\forall n, m \in \mathbb{Z}^+, n, m \geq N_k \Longrightarrow \|v_n - v_m\| < 2^{-k}.$$

Define $n_k = N_1 + \ldots + N_k$. Then $n_1 < n_2 < n_3 < \ldots$, and $\forall k \in \mathbb{Z}^+, n_k \geq N_k$. Thus, $\forall k \in \mathbb{Z}^+$, $\|v_{n_{k+1}} - v_{n_k}\| < 2^{-k}$. Thus, $\sum_{k=1}^{\infty} \left(v_{n_{k+1}} - v_{n_k}\right)$ is absolutely summable, which by assumption means that $\sum_{k=1}^{\infty} \left(v_{n_{k+1}} - v_{n_k}\right)$ is summable. Thus $\left\{\sum_{k=1}^{m} \left(v_{n_{k+1}} - v_{n_k}\right)\right\}_{m=1}^{\infty}$ converges in V. Thus,

$$\left\{ v_{n_m} = \sum_{n=1}^{m-1} \left(v_{n_{k+1}} - v_{n_k}\right) + v_{n_1} \right\}_{m=1}^{\infty}$$

converges in V. □

2.1 Operators and Functionals

> Functionals eat vectors and spit out field members.
>
> Prof. Casey Rodriguez

Let's keep an example in mind throughout this course. Let $K : [0,1] \times [0,1] \to \mathbb{C}$. Assume that K is continuous. For $f \in C[0,1]$, we can define a new function

$$Tf(x) = \int_0^1 K(x,y) f(y)\, dy.$$

Then we claim $Tf \in C[0,1]$, and $\forall \lambda_1, \lambda_2 \in \mathbb{C}, \forall f_1, f_2 \in C[0,1]$,

$$T(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 T f_1 + \lambda_2 T f_2.$$

(That is, T is linear.)

Definition 2.4. Let V, W be vector spaces. We say $T : V \to W$ is linear if $\forall \lambda_1, \lambda_2 \in \mathbb{K}, \forall v_1, v_2 \in V$,

$$T(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 T v_1 + \lambda_2 T v_2.$$

We will refer to such functions (which we commonly know as linear transformations in linear algebra) as linear operators. Recall that if $(X, d_X), (Y, d_Y)$ are metric spaces, then $T : X \to Y$ is continuous if $\forall x \in X, \forall \{x_n\} \subset X$, $x_n \to x$ implies that $T(x_n) \to T(x)$. Equivalently, $T^{-1}(U)$ is open, when U is open. We aren't quite ready to prove this yet, but the following is an important fact.

Proposition 2.5. *If $T : V \to W$ is linear, with V, W being normed vector spaces, and V is finite dimensional, then T is continuous.*

To prove this, let us consider the following.

Theorem 2.6. *A linear operator $T : V \to W$ is continuous iff there exists $C > 0$ such that $\forall v \in V$,*

$$\|Tv\|_W \leq C\|v\|_V, \tag{2.6.1}$$

where V, W are normed vector spaces.

In this case, we say T is a *bounded linear operator*. This doesn't mean that $T(V)$ is bounded. In fact, the only linear operator $T : V \to W$ for which $T(V)$ is bounded is the zero operator. This inequality says that a continuous operator $T : V \to W$ maps bounded subsets of V to bounded subsets of W.

Proof of Theorem 2.6. Assume that Eq. (2.6.1) holds. Let $v \in V$, and suppose $\{v_n\} \subset V$, $v_n \to v$. By Eq. (2.6.1),

$$0 \leq \|Tv_n - Tv\|_W = \|T(v_n - v)\|_W \leq C\|v_n - v\|_V \to 0,$$

so by the squeeze theorem, $\|Tv_n - Tv\|_W \to 0$. (I prefer an epsilon-delta argument.) For the converse, assume that T is continuous. Therefore,

$$T^{-1}(B_W(0,1)) = \{v \in V \mid Tv \in B_W(0,1)\}$$

is an open set in V. Note that since T is linear, $T(0_V) = 0_W$, and since $T^{-1}(B_W(0,1))$ is open, there exists $r > 0$ such that $B_V(0,r) \subset T^{-1}(B_W(0,1))$. Let $v \in V \setminus \{0\}$. We claim that $C = \frac{2}{r}$ is sufficient to prove our result. Consider

$$\begin{aligned}
&\left\|\frac{r}{2\|v\|_V}v\right\|_V = \frac{r}{2} < r \\
&\implies \frac{r}{2\|v\|_V}v \in B_V(0,r) \\
&\implies T\left(\frac{r}{2\|v\|_V}v\right) \in B_W(0,1) \\
&\implies \left\|T\left(\frac{r}{2\|v\|_V}\right)\right\|_W < 1 \\
&\implies \|Tv\|_W < \frac{2}{r}\|v\|_V.
\end{aligned}$$

Therefore, Eq. (2.6.1) holds with $C = \frac{2}{r}$. □

Note. We will now drop the subscripts from the norms. I.e. it should hopefully be sufficiently clear that $\|Tv\| = \|Tv\|_W$, or $\|v\| = \|v\|_V$.

Example 2.7. Consider $T : C[0,1] \to C[0,1]$, given by

$$Tf(x) = \int_0^1 K(x,y)f(y)\,dy,$$

where $K \in C([0,1] \times [0,1])$ is a bounded linear operator. Let $f \in C[0,1]$, and recall

$$\|f\|_\infty = \sup_{x \in [0,1]} |f(x)|.$$

Then for all $x \in [0,1]$,

$$\begin{aligned} |Tf(x)| &= \left| \int_0^1 K(x,y) f(y)\, dy \right| \\ &\leq \int_0^1 |K(x,y)||f(y)|\, dy \\ &\leq \int_0^1 |K(x,y)|\|f\|_\infty\, dy \\ &\leq \int_0^1 \|K\|_\infty \|f_\infty\|\, dy \\ &= \|K\|_\infty \|f\|_\infty, \end{aligned}$$

where

$$\|K\|_\infty = \sup_{\substack{x\in[0,1]\\ y\in[0,1]}} |K(x,y)|.$$

Therefore, $\|Tf\|_\infty \leq \|K\|_\infty \|f\|_\infty$.

Note. This function K is usually called the kernel of the linear operator T.

Definition 2.8. We will write $\mathcal{B}(V,W) = \{T : V \to W \mid T \text{ is a bounded linear operator}\}$. It is not hard to prove that $\mathcal{B}(V,W)$ is a vector space using properties of linear operators and continuous functions. We define the operator norm on $\mathcal{B}(V,W)$ via

$$\|T\| = \sup_{\substack{v\in V\\ \|v\|=1}} \|Tv\|.$$

Theorem 2.9. *The proposed operator norm is a norm on $\mathcal{B}(V,W)$.*

Proof. For definiteness, suppose $Tv = 0$, for all $\|v\| = 1$. Then

$$0 = T\left(\frac{v}{\|v\|}\right) = \frac{1}{\|v\|} Tv \implies Tv = 0, \quad \forall v \in V.$$

For homogeneity, consider that for any $\lambda \in \mathbb{K}$,

$$\begin{aligned} \|\lambda T\| &= \sup_{\substack{v\in V\\ \|v\|=1}} \|\lambda T v\| \\ &= \sup_{\substack{v\in V\\ \|v\|=1}} |\lambda|\|Tv\| \\ &= |\lambda| \sup_{\substack{v\in V\\ \|v\|=1}} \|Tv\| \\ &= |\lambda|\|T\|. \end{aligned}$$

For the triangle inequality, consider $S, T \in \mathcal{B}(V,W)$, and $v \in V$, $\|v\| = 1$. Then

$$\begin{aligned} \|(S+T)v\| &= \|Sv + Tv\| \\ &\leq \|Sv\| + \|Tv\| \\ &\leq \|S\| + \|T\|. \end{aligned}$$

Therefore, $\|S + T\| \leq \|S\| + \|T\|$. $\square$

Note. We would also need to show that the operator norm is well defined, i.e. that it is finite for any $T \in \mathcal{B}(V,W)$. Recall that if $T \in \mathcal{B}(V,W)$, then $\exists C > 0$ such that $\forall v \in V$, $\|Tv\| \leq C\|v\|$, therefore, $\|T\|$ is the largest of these C's for any v with unit length.

Note. If $v \neq 0$, then $\left\|T\left(\frac{v}{\|v\|}\right)\right\| \leq \|T\| \implies \|Tv\| \leq \|T\|\|v\|$.

Theorem 2.10. *If W is a Banach space, then $\mathcal{B}(V,W)$ is a Banach space.*

Proof. Suppose $\{T_n\} \subset \mathcal{B}(V,W)$ such that

$$C = \sum_{n=1}^{\infty}\|T_n\| < \infty.$$

We want to show that $\sum_{n=1}^{\infty} T_n$ is summable. Let $v \in V$, let $m \in \mathbb{Z}^+$. Then

$$\begin{aligned}\sum_{n=1}^{m}\|T_n v\| &\leq \sum_{n=1}^{m}\|T_n\|\|v\| \\ &\leq \|v\|\sum_{n=1}^{m}\|T_n\| \\ &= C\|v\|.\end{aligned}$$

Since $\{\sum_{n=1}^{m}\|T_n v\|\}_{m=1}^{\infty}$ is bounded, this implies $\sum_{n=1}^{\infty}\|T_n v\|$ converges. Thus $\sum_{n=1}^{\infty} T_n v$ is absolutely summable in W. Since W is a Banach space, $\sum_{n=1}^{\infty} T_n v$ is summable. Therefore, define $T : V \to W$ by

$$Tv = \lim_{m\to\infty}\sum_{n=1}^{m} T_n v.$$

We claim T is linear. Consider $c_1, c_2 \in \mathbb{K}$, $v_1, v_2 \in V$, and

$$\begin{aligned}T(c_1v_1 + c_2v_2) &= \lim_{m\to\infty}\sum_{n=1}^{m} T_n(c_1v_1 + c_2v_2) \\ &= \lim_{m\to\infty}\left(c_1\left(\sum_{n=1}^{m} T_n v_1\right) + c_2\left(\sum_{n=1}^{m} T_n v_2\right)\right) \\ &= c_1Tv_1 + c_2Tv_2.\end{aligned}$$

So T is linear. We will now show that T is a bounded linear operator. Let $v \in V, \|v\| = 1$. Then

$$\begin{aligned}
\|Tv\| &= \left\| \lim_{m\to\infty} \sum_{n=1}^{m} T_n v \right\| \\
&= \lim_{m\to\infty} \left\| \sum_{n=1}^{m} T_n v \right\| \\
&\leq \lim_{m\to\infty} \sum_{n=1}^{m} \|T_n v\| \\
&\leq \lim_{m\to\infty} \sum_{n=1}^{m} \|T_n\| \\
&= \sum_{n=1}^{\infty} \|T_n\| = C < \infty. \\
\implies \|Tv\| &\leq C\|v\|, \quad \forall v \in V.
\end{aligned}$$

Therefore, $T \in \mathcal{B}(V, W)$. We want to finally show that $\sum_{n=1}^{m} T_n \to T$, as $m \to \infty$. Let $v \in V$, with $\|v\| = 1$. Then

$$\begin{aligned}
\left\| Tv - \sum_{n=1}^{m} T_n v \right\| &= \left\| \lim_{m'\to\infty} \sum_{n=1}^{m'} T_n v - \sum_{n=1}^{m} T_n v \right\| \\
&= \left\| \lim_{m'\to\infty} \sum_{n=m+1}^{m'} T_n v \right\| \\
&\leq \lim_{m'\to\infty} \sum_{n=m+1}^{m'} \|T_n v\| \\
&\leq \lim_{m'\to\infty} \sum_{n=m+1}^{m'} \|T_n\| \\
&= \sum_{n=m+1}^{\infty} \|T_n\| \\
\implies \left\| T - \sum_{n=1}^{m} T_n \right\| &\leq \sum_{n=m+1}^{\infty} \|T_n\| \\
&\xrightarrow{m\to\infty} 0 \\
&\implies \sum_{n=1}^{m} T_n \to T.
\end{aligned}$$

□

Definition 2.11. If V is a normed vector space, then $V' = \mathcal{B}(V, \mathbb{K})$ is the dual space of V. Note that since $\mathbb{K}$ is always complete, V' is a Banach space.

Example 2.12. For all $p \in \mathbb{Z}^+, 1 \leq p \leq \infty$,

$$(\ell^p)' = \ell^{p'},$$

where $\frac{1}{p'} + \frac{1}{p} = 1$. For example, $(\ell^1)' = \ell^\infty$, $(\ell^2)' = \ell^2$. But $(\ell^\infty)' \neq \ell^1$.

3 Lecture 3. Baire Category Theorem and Uniform Boundedness Theorem

3.1 Subspaces and Quotient Spaces

Definition 3.1. Let V be a vector space. A subset $W \subset V$ is a subspace of V if for all $c_1, c_2 \in \mathbb{K}$, $w_1, w_2 \in W$, $c_1 w_1 + c_2 w_2 \in W$. We write in this case $W \leq V$.

Theorem 3.2. *A subspace $W \subset V$ of a Banach space is a Banach space iff $W \subset V$ is closed.*

Proof. Suppose W is a Banach space. Suppose w' is a limit point of W. Then there is a sequence $\{w_n\}$ such that $w_n \to w'$. So $\{w_n\}$ is Cauchy, and since W is Banach, $w_n \to w_0$, for some $w_0 \in W$. Since limits are unique, conclude $w' = w_0$, so $w' \in W$. In the converse direction, supposed W is closed. Then every convergent sequence in W converges to a member of W, and all convergent sequences are Cauchy, so W must be Banach. $\square$

Definition 3.3. Let $W \leq V$. We define an equivalence relation on by $v \sim v' \iff v - v' \in W$. Let us go ahead and prove $\sim$ is in fact an equivalence relation. For any $v \in V$, $v - v = 0$, which is in W, since W is a subspace, so $v \sim v$. Consider that if $v \sim v'$, then $v - v' = w$, for some $w \in W$, but recall that $-w \in W$, since W is a subspace, and hence $v' - v = -w \implies v' \sim v$. Finally, if $v \sim v'$, and $v' \sim v''$, we know $v - v' = w_1$, and $v' - v'' = w_2$, for some $w_1, w_2 \in W$. But W is a subspace, so $v'' - v = v' - w_2 - w_1 - v' = w_2 - w_1 \in W$. Thus $v \sim v''$.

Definition 3.4. Let $v \in V$. Define $[v] = \{v' \in V \mid v' \sim v\}$. Define $V/W = \{[v] \mid v \in V\}$. We call V/W the quotient space V mod W. We write $v + W = [v]$. In this case, W is necessarily a subspace of V for the following theorem to hold.

Theorem 3.5. *We claim V/W is a vector space under the operations $(v_1+W)+(v_2+W) := (v_1+v_2)+W$, for any $v_1, v_2 \in V$, and $\lambda(v + W) = \lambda v + W$, for any $v \in V, \lambda \in \mathbb{K}$.*

Note. We can identify W as $0 + W$.

Theorem 3.6. *Let $\|\cdot\|$ be a semi-norm on V. Then $E := \{v \in V \mid \|v\| = 0\}$ is a subspace of V, and $\|v + E\|_{V/E} := \|v\|, \forall v + E \in V/E$ is a norm on V/E.*

Proof. The set E is a subspace since $\forall v_1, v_2 \in E$, and $c_1, c_2 \in \mathbb{K}$,

$$\|c_1 v_1 + c_2 v_2\| \leq |c_1|\|v_1\| + |c_2|\|v_2\| = 0,$$

so $\|c_1 v_1 + c_2 v_2\| = 0 \implies c_1 v_1 + c_2 v_2 \in E$. Let us now prove that $\|\cdot\|_{V/E}$ is well-defined. Let $v, v' \in V$, with $v + E = v' + E$. Then $\exists e \in E$ such that $v = v' + e$. Then

$$\begin{aligned} \|v\| = \|v' + e\| &\leq \|v'\| + \|e\| \\ &= \|v'\|. \end{aligned}$$

This argument is symmetric in v, v', so $\|v\| = \|v'\|$, proving that $\|\cdot\|_{V/E}$ is well-defined. The remainder of the proof is standard. Most of the norm properties of $\|\cdot\|_{V/E}$ are inherited from $\|\cdot\|$. We've also equated all of the elements with zero norm in V/E by modding by E, and positive-definiteness follows from this. $\square$

3.2 Baire Category Theorem

We have one more result to talk about before moving on to some serious analysis. The following has nothing to do with category theory.

Definition 3.7. If the closure of a subset $E \subset X$ does not contain an open ball, we say that E is nowhere dense.

Theorem 3.8 (Baire's Category Theorem)**.** *If (M, d) is a complete metric space, and $\{C_n\}_{n \in \mathbb{Z}^+}$ is a collection of closed subsets of M such that $M = \bigcup_{n \in \mathbb{Z}^+} C_n$, then at least one C_n contains an open ball $B(x, r) = \{y \in M \mid d(x, y) < r\}$. More succinctly, at least one of the C_n's has an interior point.*

Baire's theorem basically states that we cannot write M as the union of nowhere dense subsets, when M is complete.

Proof. We will argue by contradiction. Suppose $\{C_n\}_{n \in \mathbb{Z}^+}$ is a collection of closed subsets of M such that $M = \bigcup_{n \in \mathbb{Z}^+} C_n$, and that each C_n is nowhere dense. We argue that M contains an open ball, since M itself is open. Therefore, $M \neq C_1$. Thus, choose $p_1 \in M \setminus C_1$. Since C_1 is closed, C_1^C is open, so there exists $\varepsilon_1 > 0$ such that $B(p_1, \varepsilon_1) \cap C_1 = \emptyset$. Now, $B(p_1, \varepsilon_1/3) \not\subset C_2$. Therefore, there exists $p_2 \in B(p_1, \varepsilon_1/3)$ such that $p_2 \notin C_2$. Since C_2 is closed, there exists $\varepsilon_2 < \varepsilon_1/3$ such that $B(p_2, \varepsilon_2) \cap C_2 = \emptyset$. Suppose there exists k points $p_1, p_2, \ldots, p_k$ and positive $\varepsilon_1, \ldots, \varepsilon_k$ such that $\varepsilon_k < \varepsilon_{k-1}/3 < \ldots < \varepsilon_1/3^{k-1}$, and $p_k \in B(p_{k-1}, \varepsilon_{k-1}/3)$, and $B(p_j, \varepsilon_k) \cap C_j = \emptyset$, for $j = 2, \ldots, k$ (*). Since $B(p_k, \varepsilon_k/3) \not\subset C_{k+1}$, there exists $p_{k+1} \in B(p_k, \varepsilon_k/3)$ such that $p_{k+1} \notin C_{k+1}$. Then there exists $\varepsilon_{k+1} < \varepsilon_k/3$ such that $B(p_{k+1}, \varepsilon_{k+1}) \cap C_{k+1} = \emptyset$. By induction, we have found a sequence of points $\{p_k\} \subset M$ and $\{\varepsilon_k\} \subset \mathbb{R}$, for all $k \in \mathbb{Z}^+$, $\varepsilon_k \in (0, \varepsilon_1)$, and (*) holds. We claim that $\{p_k\}$ is Cauchy. This follows from the fact that $\forall k, \ell \in \mathbb{Z}^+$,

$$\begin{aligned} d(p_k, p_{k+\ell}) &\leq d(p_k, p_{k+1}) + \ldots + d(p_{k+\ell-1}, p_{k+\ell}) \\ &< \frac{\varepsilon_k}{3} + \frac{\varepsilon_{k+1}}{3} + \ldots + \frac{\varepsilon_{k+\ell-1}}{3} \\ &< \frac{\varepsilon_1}{3^k} + \ldots + \frac{\varepsilon_1}{3^{k+\ell}} \\ &< \varepsilon_1 \sum_{n=k}^{\infty} \frac{1}{3^n} \\ &= \frac{\varepsilon_1}{3^k} \frac{1}{1 - \frac{1}{3}} \\ &= \frac{\varepsilon_1}{2} 3^{-k+1}. \end{aligned}$$

Since this quantity is independent of ℓ, and grows small with as k grows large, we can say $\{p_k\}$ is

Cauchy. Since M is complete, there is some $p \in M$ such that $p_k \to p$. Now for all $k \in \mathbb{Z}^+$,

$$
\begin{aligned}
d(p_{k+1}, p_{k+1+\ell}) &< \varepsilon_{k+1}\left(\frac{1}{3} + \frac{1}{3^2} + \ldots + \frac{1}{3^\ell}\right) \\
&< \varepsilon_{k+1}\sum_{m=0}^{\infty} 3^{-m} \\
&= \varepsilon_{k+1}\frac{3}{2} \\
\implies d(p_{k+1}, p) &\leq \frac{3}{2}\varepsilon_{k+1} < \frac{1}{2}\varepsilon_k && \text{(as } \ell \to \infty\text{)} \\
\implies d(p_k, p) &\leq d(p_k, p_{k+1}) + d(p_{k+1}, p) \\
&\leq \frac{1}{3}\varepsilon_k + \frac{1}{2}\varepsilon_k < \varepsilon_k \\
&\implies p \in B(p_k, \varepsilon_k) \\
&\implies p \notin C_k.
\end{aligned}
$$

Since k is arbitrary, $p \notin \bigcup_{k\in\mathbb{Z}^+} C_k = M$, which is a contradiction. □

3.3 Uniform Boundedness Theorem

Let's use our powerful new result to prove some fundamental results of functional analysis.

Theorem 3.9 (Uniform Boundedness Theorem)**.** *Let B be a Banach space. Let $\{T_n\}$ be a sequence in $\mathcal{B}(B, V)$, where V is a normed vector space. If $\forall b \in B$,*

$$\sup_{n\in\mathbb{Z}^+} \|T_n b\| \leq \infty,$$

then

$$\sup_{n\in\mathbb{Z}^+} \|T_n\| < \infty.$$

Proof. For all $k \in \mathbb{Z}^+$, define

$$C_k = \{b \in B \mid \|b\| < 1 \text{ and } \sup_{n\in\mathbb{Z}^+} \|T_n b\| \leq k\}.$$

We claim C_k is closed. If $\{b_n\} \subset C_k$, and $b_n \to b$, then $\|b\| = \lim_{n\to\infty}\|b_n\| \leq 1$. Furthermore, for all $m \in \mathbb{Z}^+$, $\|T_m b\| = \lim_{n\to\infty}\|T_m b_n\| \leq k$. Thus, $b \in C_k$, which implies C_k is closed. Now,

$$E := \{b \in B \mid \|b\| \leq 1\} = \bigcup_{k\in\mathbb{Z}^+} C_k.$$

Using a similar argument as above, we see that E is closed, thus Banach, since it is a closed subset of a Banach space. Therefore, by Baire's theorem, there exists $k \in \mathbb{Z}^+$ such that C_k contains an open ball $B(b_0, \delta_0)$ for some $b_0 \in C_k$, $\delta_0 > 0$. If $b \in B(0, \delta_0)$ (that is, $\|b\| < \delta_0$), then $b_0 + b \in B(b_0, \delta_0)$. Thus,

$$\sup_{n\in\mathbb{Z}^+} \|T_n(b_0 + b)\| \leq k.$$

Then

$$\begin{aligned}\sup_{n\in\mathbb{Z}^+}\|T_n b\| &\le \sup_{n\in\mathbb{Z}^+}\|-T_n b_0 + T_n(b_0+b)\| \\ &\le \sup_{n\in\mathbb{Z}^+}\|T_n b_0\| + \sup_{n\in\mathbb{Z}^+}\|T_n(b_0+b)\| \\ &\le k+k = 2k.\end{aligned}$$

Let $n \in \mathbb{Z}^+$, and let $\|b\| = 1$. Then

$$\begin{aligned}\left\|T\left(\frac{\delta_0}{2}b\right)\right\| &\le 2k \\ \implies \|T_n b\| &\le \frac{4k}{\delta_0} \\ \implies \|T_n\| &\le \frac{4k}{\delta_0} \\ \implies \sup_{n\in\mathbb{Z}^+}\|T_n\| &\le \frac{4k}{\delta_0}.\end{aligned}$$

□

4 Lecture 4. Open Mapping Theorem and Closed Graph Theorem

Let's recap what we talked about last time.

Theorem 4.1 (Baire's Category Theorem)**.** *If M is a complete metric space, and $\{C_k\}$ is a collection of closed sets and*

$$M = \bigcup_{k\in\mathbb{Z}^+} C_k,$$

then at least one of C_k contains an interior point.

4.1 Open Mapping Theorem

Let's talk about another big-name theorem.

Theorem 4.2 (Open Mapping Theorem)**.** *If B_1, B_2 are Banach spaces, and $T \in \mathcal{B}(B_1, B_2)$ is surjective, then T is an open mapping, that is, for every $U \subset B_1$, U open implies $T(U)$ is open.*

Proof. We will first show this result holds for a certain class of open set in B_1, and we will generalize using properties of linear maps from B_1 to B_2. Let $T \in \mathcal{B}(B_1, B_2)$ be a surjective map. Consider $B(0,1)$. We want to show that $T(B(0,1))$ contains an open ball centered at 0. Since T is surjective, we can write

$$B_2 = \bigcup_{n\in\mathbb{Z}^+} \overline{T(B(0,n))}.$$

Then by Baire's theorem, there is some $n_0 \in \mathbb{Z}^+$ such that $\overline{T(B(0,n_0))}$ contains an open ball, which by the linearity of T, implies that $n_0\overline{T(B(0,1))} = \{n_0 x \mid x \in \overline{T(B(0,1))}\}$ contains an open ball. This implies that $\overline{T(B(0,1))}$ contains an open ball, by a scaling argument. Thus, there exists $v_0 \in B_2$ and $r > 0$ such that $B(v_0, 4r) \subset \overline{T(B(0,1))}$. Thus, there exists $v_1 = T(u_1) \in T(B(0,1))$, for some $u_1 \in B_1$ such that $\|v_0 - v_1\| < 2r$. Thus,

$$B(v_1, 2r) \subset B(v_0, 4r) \subset \overline{T(B(0,1))}.$$

Let $v \in V$, such that $\|v\| < r$. Then

$$\begin{aligned}
\frac{1}{2}(2v + v_1) &\in \frac{1}{2}\overline{T(B(0,1))} \\
&= \overline{T(B(0,1/2))} \\
\implies v &= -T\left(\frac{u_1}{2}\right) + \frac{1}{2}(2v + v_1) \\
&\in \overline{\overbrace{-T\left(\frac{u_1}{2}\right)}^{\text{scalar}} + \overline{T(B(0,1/2))}} \\
&= T\left(\underbrace{-u_1/2 + B(0,1/2)}_{\subset B(0,1)}\right) \subset \overline{T(B(0,1))}.
\end{aligned}$$

By a scaling argument, for $n \in \mathbb{Z}^+$,

$$\begin{aligned}
B(0,r) &\subset \overline{T(B(0,1))} \\
\implies B(0, 2^{-n}r) &= 2^{-n}B(0,r) \\
&\subset 2^{-n}\overline{T(B(0,1))} = \overline{T(B(0,2^{-n}))}.
\end{aligned}$$

Now, we will prove that $B(0,r/2) \subset T(B(0,1))$. Let $v \in V$, $\|v\| < r/2$. Then $v \in \overline{T(B(0,1/2))}$, which implies that there is some $b_1 \in B(0,1/2)$ such that $\|v - Tb_1\| < r/4$. Thus $v - Tb_1 \in \overline{T(B(0,1/4))}$, which implies that there exists $b_2 \in B(0,1/4)$ such that $\|v - Tb_1 - Tb_2\| < r/8$. Continuing inductively, we obtain a sequence $\{b_k\}$ in B_1 such that

1. $\|b_k\| < 2^{-k}$
2. $\|v - \sum_{k=1}^{\infty} Tb_k\| < 2^{(-n-1)}r$.

We claim the series $\sum_{k=1}^{\infty} b_k$ is absolutely summable, which implies that there exists $b \in B_1$ such that $b = \sum_{k=1}^{\infty} b_k$. Moreover,

$$\|b\| = \lim_{n\to\infty}\left\|\sum_{k=1}^{n} b_k\right\| \le \lim_{n\to\infty}\sum_{k=1}^{n}\|b_k\| \le \sum_{k=1}^{\infty}\|b_k\| < \sum_{k=1}^{\infty} 2^{-k} = 1.$$

Moreover, since T is continuous,

$$\begin{aligned}
Tb = \lim_{n\to\infty} T\left(\sum_{k=1}^{n} b_k\right) &= \lim_{n\to\infty}\sum_{k=1}^{n} Tb_k = v \\
&\implies v \in T(B(0,1)).
\end{aligned}$$

Thus, $B(0,r/2) \subset T(B(0,1))$. Now, suppose $U \subset B_1$ is open, and let $b_1 \in U$, and $b_2 \in B_2$, such that $b_2 = Tb_1 \in T(U)$. Then $\exists \varepsilon > 0$ such that $b_1 + B(0,\varepsilon) = B(b_1,\varepsilon) \subset U$. Since $\exists \delta > 0$ such that

$$B(0,\delta) \subset T(B(0,1))$$

we have

$$\begin{aligned}
\implies B(b_2, \varepsilon\delta) &= b_2 + \varepsilon B(0, \delta) \\
&\subset b_2 + \varepsilon T(B(0,1)) \\
&= T(b_1) + \varepsilon T(B(0,1)) \\
&= T(b_1 + \varepsilon B(0,1)) \\
&= T(b_1 + B(0, \varepsilon)) \\
&\subset T(U).
\end{aligned}$$

□

Corollary 4.2.1. *If B_1, B_2 are Banach spaces, and $T \in \mathcal{B}(B_1, B_2)$ is bijective, then $T^{-1} \in \mathcal{B}(B_2, B_1)$.*

Proof. Note that T^{-1} is continuous iff $\forall U \subset B_1$, U open, $(T^{-1})^{-1}(U) = T(U)$ is open, which is true by the open mapping theorem, since T is surjective. It remains to show that T is linear, which is left as an exercise. □

4.2 Closed Graph Theorem

From the open mapping theorem, we get the closed graph theorem. Before we state this important theorem, we need an intermediate result.

Theorem 4.3. *If B_1, B_2 are Banach spaces, then $B_1 \times B_2$, under the norm*

$$\|(b_1, b_2)\| := \|b_1\| + \|b_2\|, \quad b_1 \in B_1, b_2 \in B_2,$$

is Banach.

The proof for this result is standard.

Theorem 4.4 (Closed Graph Theorem)**.** *If B_1, B_2 are Banach spaces, and $T : B_1 \to B_2$ is linear, then $T \in \mathcal{B}(B_1, B_2)$ iff*

$$\Gamma(T) = \{(u, Tu) \mid u \in B_1\}$$

is closed.

Proof. For the forward direction, let $T \in \mathcal{B}(B_1, B_2)$. Let $\{(u_n, Tu_n)\}$ be a sequence in $\Gamma(T)$ such that $u_n \to u$, $Tu_n \to v$, for some $u \in B_1, v \in B_2$. Then

$$v = \lim_{n\to\infty} Tu_n = T\left(\lim_{n\to\infty} u_n\right) = Tu.$$

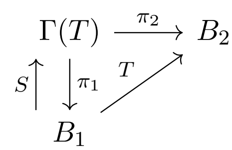
Thus, $(u, v) = (u, Tu) \in \Gamma(T)$. For the opposite direction, assume $\Gamma(T)$ is closed, and define

1. $\pi_1 : \Gamma(T) \to B_1$ by $\pi_1(u, Tu) = u$,
2. $\pi_2 : \Gamma(T) \to B_2$ by $\pi_2(u, Tu) = Tu$.

Note that $\Gamma(T)$ is a Banach space. Moreover, $\pi_1 \in \mathcal{B}(\Gamma(T), B_1)$, and $\pi_2 \in \mathcal{B}(\Gamma(T), B_2)$. To show this, note that

$$\|\pi_2(u, v)\| = \|v\| \leq \|u\| + \|v\| = \|(u, v)\|,$$

and the same follows for π_1. Moreover moreover, $\pi_1 : \Gamma(T) \to B_1$ is bijective. Thus, $S := \pi_1^{-1} : B_1 \to \Gamma(T)$ is a bounded linear operator by Corollary 4.2.1, which implies that $T = \pi_2 \circ S : B_1 \to B_2$.



Therefore, $T \in \mathcal{B}(B_1, B_2)$. $\square$

4.3 Hahn-Banach Theorem Intro

The Hahn-Banach theorem tries to answer the following question.

Q: Given a general nontivial normed space V, does $V' = \{0\}$?

Example 4.5. We stated earlier but did not prove that $(\ell^p)' = l^{p'}$, where $\frac{1}{p} + \frac{1}{p'} = 1, 1 \leq p < \infty$, and that $(c_0)' = \ell^1$ (the set of sequences which converge to 0).

To answer this question, we will need a certain axiom from set theory.

Definition 4.6. A partial order on a set E is a relation $\preceq$ on E such that

1. $\forall e \in E, e \preceq e$,
2. $\forall e, f \in E$, $e \preceq f$ and $f \preceq e \Longrightarrow e = f$,
3. $\forall e, f, g \in E$, $e \preceq f$ and $f \preceq g \Longrightarrow e \preceq g$.

An upper bound of a set $D \subset E$ is an element $e \in E$ such that $\forall d \in D, d \preceq e$. A maximal element of E is an element $e \in E$ such that if $f \in E$, and $e \preceq f$, then $e = f$. Similarly defined is minimal element of E.

Definition 4.7. If $(E, \preceq)$ is a partially ordered set (poset), a subset $C \subset E$ is a chain in E if $\forall e, f \in C$, either $e \preceq f$, or $f \preceq e$.

The following is an axiom of Zermelo-Fraenkel set theory.

Lemma 4.8 (Zorn's Lemma)**.** *If every chain in a nonempty partially ordered set E has an upper bound, then E has a maximal element.*

Definition 4.9. A Hamel basis $H \subset V$, where V is a vector space, is a linearly independent set such that every element of V is a finite linear combination of elements of H.

5 Lecture 5. Hahn-Banach Theorem

5.1 Proof of Hahn-Banach Theorem

We continue with the statement and proof of the Hahn-Banach theorem.

Example 5.1. The set $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ is a Hamel basis for $\mathbb{R}^2$.

Before we get our heads wrapped around Hahn-Banach, we should get familiar with applying Zorn's lemma. Linear algebra is usually treated in the finite-dimensional setting, and we understand that finite-dimensional vector spaces have bases. Since we're in functional analysis, we should study whether or not infinite-dimensional vector spaces have Hamel bases.

Theorem 5.2. *If V is a vector space, then V has a Hamel basis.*

Proof. Let E be the set of all linearly independent subsets of V. Define a partial ordering $\preceq$ on E by inclusion: for $e_1, e_2 \in V$, $e_1 \preceq e_2 \iff e_1 \subset e_2$. Let C be a chain in E. Define $c = \bigcup_{e \in C} e$. We claim that $c \in E$, that is c is a linearly independent subset of V. To show this, let $v_1, \ldots, v_N \in c$. Then there exists $e_1, e_2, \ldots, e_N \in C$ such that $\forall j \in [N], v_j \in e_j$. Since C is a chain, $\exists J$ such that $\forall j \in [N]$, $e_j \preceq e_J$, that is, $e_j \subset e_J$. Therefore, $v_1, \ldots, v_N \in e_J$. Thus, $v_1, \ldots, v_N$ are linearly independent, since $e_J \in E$. Thus every finite collection of vectors in c is linearly independent, thus, $c \in E$. We've shown every chain in E has an upper bound, so we conclude that E has a maximal element H. We claim that H spans V. By contradiction, assume that there is some $v \in V$ such that v cannot be written as a finite linear combination of vectors in H. But then $H \cup \{v\}$ is linearly independent, and also $H \subset H \cup \{v\}$, which is a contradiction. Thus, H is a Hamel basis of V. □

Let's state the Hahn-Banach theorem.

Theorem 5.3 (Hahn-Banach Theorem)**.** *Let V be a normed space. Let $M \leq V$, and let $u : M \to \mathbb{C}$ be a linear functional satisfying*

$$\forall t \in M, \quad |u(t)| \leq C\|t\| \tag{5.1.1}$$

Then $\exists U \in V'$ such that $U|_M = u$, and

$$\forall t \in V, \quad |U(t)| \leq C\|t\|.$$

We say that U is a continuous extension of u to V. Let us state a lemma that will help us prove the Hahn-Banach theorem.

Lemma 5.4. *If V is a normed vector space, and $M \leq V$, and $u : M \to \mathbb{C}$ is linear such that*

$$\forall t \in M, \quad |u(t)| \leq C\|t\|,$$

and $x \notin M$, then $\exists u' : M_0 \to \mathbb{C}$ which is linear, where $M_0 = \{t + ax \mid t \in M, a \in \mathbb{C}\}$, and $u'|_M = u$, and

$$\forall t' \in M_0, \quad |u'(t')| \leq C\|t'\|.$$

Proof. First note that if $t' \in M_0$, then $\exists! t \in M$ and $a \in \mathbb{C}$ such that $t' = t + ax$. By contradiction, if $t + ax = \tilde{t} + \tilde{a}x$, then $(a - \tilde{a})x = \tilde{t} - t$. But this implies that $x \in M$. Also, $a = \tilde{a}$, iff $\tilde{t} = t$. Thus, upon choosing $\lambda \in \mathbb{C}$, the map

$$u'(t + ax) = u(t) + a\lambda, \quad \forall t \in M$$

is well-defined on M_0. Also, $u' : M_0 \to \mathbb{C}$ is linear. WLOG we can assume $C = 1$. (Why?) We want to choose $\lambda \in \mathbb{C}$ such that

$$\forall t \in M, \forall a \in \mathbb{C}, \quad |u(t) + a\lambda| \leq \|t + ax\|.$$

When λ is chosen, u' is our desired extension. In the case that $a = 0$, the above inequality holds, so assume that $a \neq 0$. Let $t \in M$. Consider

$$\begin{aligned} \frac{|u(t) + a\lambda|}{|a|} &\leq \frac{\|t + ax\|}{|a|} \\ \iff \left| u\left(\frac{t}{a}\right) - \lambda \right| &\leq \left\| \frac{t}{a} - x \right\| \\ \iff |u(t) - \lambda| &\leq \|t - x\|. \end{aligned}$$

We will find a suitable λ by choosing a suitable $\operatorname{Re}\lambda$ and $\operatorname{Im}\lambda$. First, we claim that there exists $\alpha \in \mathbb{R}$ satisfying

$$|\operatorname{Re} u(t) - \alpha| \leq \|t - \alpha\|,$$

Note that

$$|\operatorname{Re} u(t)| \leq |u(t)| \leq \|t\|.$$

Also, $\forall t_1, t_2 \in M$,

$$\begin{aligned} \operatorname{Re} u(t_1) - \operatorname{Re} u(t_2) &= \operatorname{Re} u(t_1 - t_2) \\ &\leq |\operatorname{Re} u(t_1 - t_2)| \\ &\leq \|t_1 - t_2\| \\ &\leq \|t_1 - x\| + \|t_2 - x\| \\ \implies \operatorname{Re} u(t_1) - \|t_1 - x\| &\leq \operatorname{Re} u(t_2) + \|t_2 - x\|, \quad \forall t_1, t_2 \in M \\ \implies \sup_{t \in M}\{\operatorname{Re} u(t_1) - \|t - x\|\} &\leq \operatorname{Re} u(t_2) + \|t_2 - x\|, \quad \forall t_2 \in M \\ \implies \sup_{t \in M}\{\operatorname{Re} u(t) - \|t - x\|\} &\leq \inf_{t \in M}\{\operatorname{Re} u(t) + \|t - x\|\}. \end{aligned}$$

We can choose an α between the LHS and RHS of the above inequality. Consider that $\forall t \in M$,

$$\begin{aligned} \operatorname{Re} u(t) - \|t - x\| \leq \alpha &\leq \operatorname{Re} u(t) + \|t - x\| \\ \implies -\|t - x\| &\leq \alpha - \operatorname{Re} u(t) \leq \|t - x\| \\ \implies |\operatorname{Re} u(t) - \alpha| &\leq \|t - x\|. \end{aligned}$$

We can repeat this argument for the imaginary part of λ, completing our proof. □

Proof of Theorem 5.3. Let

$$E = \{(\mu, N) \mid N \leq V, M \subset N, \mu \text{ is a continuous extension of } u \text{ to } N\}.$$

Note that E is nonempty, since $(u, M) \in E$. Define $\preceq$ on E such that for $(\mu_1, N_1), (\mu_2, N_2) \in E$, $(\mu_1, N_1) \preceq (\mu_2, N_2)$ iff $N_1 \subset N_2$ and $\mu_2|_{N_1} = \mu_1$. We claim that $\preceq$ is a partial ordering on E. Let $C = \{(\mu_i, N_i) \mid i \in I\}$ be a chain in E, for some index set I. Then $\forall i, j \in I$, either $(\mu_i, N_i) \preceq (\mu_j, N_j)$, or vice versa. Let

$$N = \bigcup_{i \in I} N_i.$$

We claim that N is a subspace. Let $v_1, v_2 \in N$, and $a_1, a_2 \in \mathbb{C}$. Then $\exists i, j \in I$ such that $v_i \in N_i$, and $v_j \in N_j$. Then since C is a chain, WLOG $N_i \subset N_j$. Thus, $v_1, v_2 \in N_j$. Thus, $a_1 v_1 + a_2 v_2 \in N_j$, which is a subset of N, thus N is a subspace of V. Define $\mu : N \to \mathbb{C}$ by $\mu(t) = \mu_i(t)$ when $t \in N_i$. To check that this mapping is well defined, consider that if $t \in N_i \cap N_j$, and since C is a chain, WLOG assume that $(\mu_i, N_i) \preceq (\mu_j, N_j)$. Then $\mu_j|_{N_i} = \mu_i$, so $\mu_i(t) = \mu_j(t)$, showing μ is well-defined. It is not hard to prove that μ is linear, and $\mu \in N'$, and a continuous extension of all the μ_i's (by definition). Thus, $\forall i \in I$, $(\mu_i, N_i) \preceq (\mu, N)$, so (μ, N) is an upper bound of C. Then by Zorn's lemma, E has a maximal element (U, N_{max}). We claim $N_{\text{max}} = V$. By contradiction, let $x \notin N_{\text{max}}$. By Lemma 5.4, there exists a continuous extension ν of U to $(N_{\text{max}} + \mathbb{C}x) := \{t + ax \mid t \in N_{\text{max}}, a \in \mathbb{C}\}$. Clearly $(U, N_{\text{max}}) \preceq (\nu, N_{\text{max}})$, which is a contradiction. Therefore, U is a continuous extension of u to V. □

6 Lecture 6. Double Dual and Outer Measure of Subsets of Reals

6.1 Applying Hahn-Banach

Theorem 6.1. *If V is a normed vector space, then $\forall v \in V \setminus \{0\}$, $\exists f \in V'$ such that*

$$\|f\| = 1, \quad f(v) = \|v\|.$$

Proof. Define $u : \text{span}(v) \to \mathbb{C}$ by $u(\lambda v) = \lambda \|v\|, \forall \lambda \in \mathbb{C}$. Then $|u(t)| \leq \|t\|, \forall t \in \text{span}(v)$, and $u(v) = \|v\|$. Then, by Hahn-Banach, there exists $f \in V'$ such that

$$|f(t)| \leq \|t\|, \forall \quad t \in V.$$

Then $f(v) = u(v) = \|v\|$, which implies. Since $|f(t)| \leq \|t\|, \forall t \in V$, $\|f\| \leq 1$. But

$$1 = f\left(\frac{v}{\|v\|}\right) \leq \|f\| \implies \|f\| = 1.$$

□

6.2 Double Dual

Definition 6.2. The double dual of V is $V'' := (V')'$.

Example 6.3. Let $v \in V$. Define $T_v : V' \to \mathbb{C}$ by by $T_v(v') = v'(v), \forall v' \in V'$. Then $T_v \in V''$. Linearity is clear. We claim T_v is bounded, since

$$|T_v(v')| = |v'(v)| \leq \|v'\| \|v\|$$
$$\implies T_v \in (V')' = V'' \text{ and } \|T_v\| \leq \|v\|.$$

Definition 6.4. If V, W are normed spaces, then $T \in \mathcal{B}(V, W)$ is an isometry or isometric if

$$\|Tv\| = \|v\|, \quad \forall v \in V.$$

Theorem 6.5. *Let $v \in V$. Define $T_v : V' \to \mathbb{C}$ by $T_v(v') = v'(v), \forall v' \in V'$. Then $T : V \to V''$, where $v \mapsto T_v$, is isometric.*

Proof. Let $v \in V$. We've shown already that $v \mapsto T_v \in \mathcal{B}(V, W)$, and $\|T_v\| \leq \|v\|$. We only need to show that $\|v\| \leq \|T_v\|$. This is clear if $v = 0$, so suppose $v \neq 0$. Then $\exists f \in V'$ such that $\|f\| = 1$, and $f(v) = \|v\|$. Then

$$\|v\| = |f(v)| \leq \|T_v\| \|f\| = \|T_v\|,$$

so $\|T_v\| = \|v\|$. □

Definition 6.6. A Banach space V is reflexive if $V = V''$ in the sense that $v \mapsto T_v$ is surjective.

Example 6.7. For $1 < p < \infty$, ℓ^p is reflexive. Recall that $(\ell^1)' = \ell^\infty$, but $(\ell^\infty)' \neq \ell^1$, so ℓ^1 is not reflexive.

Example 6.8. The space c_0 of sequences converging to 0 is not reflexive. The dual $(c_0)' = \ell^1$, and $(\ell^1)' = \ell^\infty$, which is not equal to c_0.

6.3 Lebesgue Measure and Integration

6.3.1 Why Lebesgue Integration?

Why not stick to Riemann integration?

1. Lebesgue integration has much better convergence theorems. Riemann integration has the constrictive theorem that the uniform limit of Riemann integrable functions is Riemann integrable, and that in uniformly convergent sequences of functions, we can swap the order of integration and limit. But there are much more useful theorems we can use when treating Lebesgue integration.

2. Consider the space of Riemann integrable functions on $[a, b]$, for example, $a = 0, b = 1$. Then consider
$$L^1_R([0,1]) = \{f : [0,1] \to \mathbb{C} \mid f \text{ is Riemann integrable on } [0,1]\}.$$
Define
$$\|f\|_1 = \int_0^1 |f(x)|\, dx, \quad \forall f \in L^1_R([0,1]).$$
(Note that this is a seminorm, but imagine that it is a norm.) Then $L^1_R([0,1])$ is not Banach. In functional analysis, we are interested in complete spaces.

When treating integration in $\mathbb{R}$, we should think of integration as the theory of "finding area under the curve." Consider the function
$$1_E(x) = \begin{cases} 1 & x \in E \\ 0 & x \in E \end{cases},$$
for some set $E \subset \mathbb{R}$. If $E = [a, b]$, then
$$\int 1_{[a,b]}(x)\, dx = b - a.$$
For more general E,
$$\int 1_E(x)\, dx = m(E),$$
where $m(E)$ is some notion of measure, or length of E. Our task now is to find out what we mean by $m(E)$. We'd like the following properties to be satisfied by this function m:

1. $m(E)$ is defined $\forall E \subset \mathbb{R}$.

2. If I is an interval, then $m(I) = \ell(I)$, that is, the length of I.

3. If $\{E_n\}$ is a countable collection of disjoint sets, then $m(\cup_n E_n) = \sum_n m(E_n)$.

4. We would like m to be translation invariant. If $E \subset \mathbb{R}$, then $m(E) = m(\{x + e \mid e \in E\})$, for $x \in \mathbb{R}$.

Unfortunately. This is impossible. Such a function $m : 2^{\mathbb{R}} \to [0, \infty)$ does not exist. If you assume that m exists satisfying the four properties above, then you can find a subset of $\mathbb{R}$ having infinite measure, and simultaneously measure 0. We instead can drop the requirement that m is defined over all sets $E \subset \mathbb{R}$, and instead try to find an m that is defined satisfying $2, 3, 4$, on a sufficiently large subset of $2^{\mathbb{R}}$. The ultimate measure we are interested in is the Lebesgue measure, and this subset of $2^{\mathbb{R}}$ is the set of Lebesgue measurable sets.

6.3.2 Outer Measure, Caratheodory

Let us construct the Lebesgue measure. This construction is due to Caratheodory. We will begin by constructing $m^* : 2^{\mathbb{R}} \to [0, \infty)$ satisfying constraint 2 (above), 4, and almost 3. We will restrict m^* to certain well-behaved subsets of $\mathbb{R}$, which will yield our desired m. This m^* we are interested in is called the outer measure.

Note. If $I \subset \mathbb{R}$ is an interval, then we will write $\ell(I)$ to be its length.

Definition 6.9. For $A \subset \mathbb{R}$, we define the outer measure of A

$$m^*(A) = \inf \left\{ \sum_n \ell(I_n) \mid \{I_n\} \text{ is a countable collection of open intervals such that } A \subset \bigcup_n I_n \right\}$$

Example 6.10. The set $\{1\}$ has outer measure 0. For any $\varepsilon > 0$, the interval $(1 - \varepsilon/2, 1 + \varepsilon/2)$ is an open interval cover of $\{1\}$. Thus,

$$m^*(\{0\}) \leq \varepsilon \xrightarrow{\varepsilon \to 0} m^*(\{0\}) = 0.$$

We can use a similar argument to prove that the outer measure of any countable set is 0.

Theorem 6.11. *If $A \subset \mathbb{R}$ is countable, then $m^*(A) = 0$.*

Proof. If A is countable infinite, then $A = \{a_n \mid n \in \mathbb{Z}^+\}$. Let $\varepsilon > 0$. We will show that $m^*(A) \leq \varepsilon$. For each $n \in \mathbb{Z}^+$, let $I_n = (a_n - \varepsilon/2^{n+1}, a_n + \varepsilon/2^{n+1})$. Then $\{I_n\}$ is an open interval cover of A. Thus,

$$\begin{aligned} m^*(A) &\leq \sum_{n=1}^{\infty} \ell(I_n) \\ &= \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon. \end{aligned}$$

Thus, $m^*(A) = 0$. □

We can generalize this argument to prove that m^* almost satisfies 3.

Theorem 6.12. *If $A \subset B$, then $m^*(A) \leq m^*(B)$.*

Proof. Any open interval covering of B is an open interval covering of A. □

Theorem 6.13. *Let $\{A_n\}$ be a countable collection of subsets of $\mathbb{R}$. Then $m^*(\cup_n A_n) \leq \sum_n m^*(A_n)$.*

Proof. If $\exists n \in \mathbb{Z}^+$ such that $m^*(A_n) = \infty$, or $\sum_n m^*(A_n) = \infty$, then our result is true. Suppose that $\forall n \in \mathbb{Z}^+$, $m^*(A_n) < \infty$, and $\sum_n m^*(A_n) < \infty$. Let $\varepsilon > 0$. For each $n \in \mathbb{Z}^+$, let $\{I_{n_k}\}_{k\in\mathbb{Z}^+}$ be a collection of open intervals such that

$$A_n \subset \bigcup_{k\in\mathbb{Z}^+} I_{n_k},$$

and

$$\sum_{k=1}^{\infty} \ell(I_{n_k}) < m^*(A_n) + \frac{\varepsilon}{2^n}.$$

Then

$$\begin{aligned}
\bigcup_{n\in\mathbb{Z}^+} A_n &\subset \bigcup_{n\in\mathbb{Z}^+, k\in\mathbb{Z}^+} I_{n_k} \\
\implies m^*(A_n) &\leq \sum_{n,k} \ell(I_{n_k}) \\
&= \sum_n \sum_k \ell(I_{n_k}) \\
&< \sum_n \left(m^*(A_n) + \frac{\varepsilon}{2^n} \right) \\
&= \varepsilon + \sum_n m^*(A_n).
\end{aligned}$$

Our result is proven by taking $\varepsilon \to 0$. □

That the outer measure m^* satisfies condition 4 is left as an exercise.

7 Lecture 7. Sigma Algebras

Theorem 7.1. *If $I \subset \mathbb{R}$ is an interval, then $m^*(I) = \ell(I)$.*

Proof. Suppose $I = [a, b]$. Then $I \subset (a - \varepsilon, b + \varepsilon)$, $\forall \varepsilon > 0$. Thus, $m^*(I) \leq \ell((a - \varepsilon, b + \varepsilon)) = b - a + 2\varepsilon, \forall \varepsilon > 0$. Since ε is arbitrary, $m^*(I) \leq b - a$. Now, we will show that $b - a \leq m^*(I)$. Let $\{I_n\}$ be a collection of open intervals covering I. By Heine-Borel, $[a, b]$ is compact, so let $I_1, \ldots, I_N$ be a finite subcovering of I. Since $a \in \cup_{j=1}^N I_j, \exists k_1$ such that $a \in I_{k_1}$. Rearrange indices, and let $a \in I_1$. Write $I_1 = (a_1, b_1)$. If $b_1 < b$, then $b_1 \in [a, b] \implies \exists k_2, b_1 \in I_{k_2}$, and by rearranging indices, we can write $b_1 \in I_2$. Continuing in this manner, we can conclude that there exists K, such that $1 \leq K \leq N$, and $\forall k = 1, \ldots, K - 1$, $b_k \leq b$, and $a_{k+1} < b_k < b_{k+1}$, with $b < b_K$. We justify this last inequality by the fact that $I_1, \ldots, I_N$ is a finite list. Then

$$\begin{aligned}
\sum_{n=1}^{\infty} \ell(I_n) &\geq \sum_{k=1}^{N} \ell(I_k) \\
&\geq \sum_{k=1}^{K} \ell(I_k) \\
&= (b_K - a_K) + (b_{K-1} - a_{K-1}) + \ldots + (b_1 - a_1) \\
&= b_K + (b_{K-1} - a_K) + (b_{K-2} - a_{K-1}) + \ldots + (b_1 - a_2) - a_1 \\
&\geq b_K - a_1 \\
&\geq b - a.
\end{aligned}$$

Thus, $m^*(I) \geq b - a$, so $m^*(I) = b - a$. If I is any finite interval $[a, b], (a, b], [a, b), (a, b)$, then $\forall \varepsilon > 0$, $[a + \varepsilon, b - \varepsilon] \subset I \subset [a - \varepsilon, b + \varepsilon]$, and using the squeeze theorem, we can conclude that $m^*(I) = b - a$. □

Exercise 7.1. If $I = \mathbb{R}, (-\infty, a), (a, \infty), (-\infty, a]$, or $[a, \infty)$, then prove that $m^*(I) = \infty$.

Theorem 7.2. *For all $A \subset \mathbb{R}$, and $\varepsilon > 0$, there exists an open set O, such that $A \subset O$, and $m^*(A) \leq m^*(O) \leq m^*(A) + \varepsilon$.*

Proof. If $m^*(A) = \infty$, then this is clear using the extended real number line. Suppose $m^*(A) < \infty$. Let $\varepsilon > 0$ be given, and let $\{I_n\}$ be an open interval cover of A, with $\sum_n \ell(I_n) \leq m^*(A) + \varepsilon$. Consider $O := \bigcup_n I_n$. Recall that O is open, since it is the union of open sets, and consider that $A \subset O$. Thus,

$$m^*(O) = m^*\left(\cup_n I_n\right) \leq \sum_n m^*(I_n) \leq \sum_n \ell(I_n) \leq m^*(A) + \varepsilon.$$

□

7.1 Lebesgue Measurable Sets

Definition 7.3. A set $E \subset \mathbb{R}$ is Lebesgue measurable if $\forall A \subset \mathbb{R}$, $m^*(A) = m^*(A \cap E) + m^*(A \cap E^C)$.

Some remarks.

1. Since $\forall A, E$, $A = (A \cap E) \cup (A \cap E^C)$, we have

$$m^*(A) \leq m^*(A \cap E) + m^*(A \cap E^C).$$

Thus, E is Lebesgue measurable if $\forall A \subset \mathbb{R}$, $m^*(A) \geq m^*(A \cap E) + m^*(A \cap E^C)$.

Theorem 7.4. *The empty set $\emptyset$ is measurable, $\mathbb{R}$ is measurable, and $E \subset \mathbb{R}$ is measurable iff E^C is measurable.*

Theorem 7.5. *If $m^*(E) = 0$, then E is measurable.*

Proof. Let $A \subset \mathbb{R}$. Then $A \cap E \subset E$. Thus

$$m^*(A \cap E) \leq m^*(E) = 0 \Longrightarrow m^*(A \cap E) = 0.$$

Therefore,

$$m^*(A \cap E) + m^*(A \cap E^C) = m^*(A \cap E^C) \leq m^*(A).$$

Thus, E is measurable. □

Theorem 7.6. *If $E_1, E_2 \subset \mathbb{R}$ are measurable, then $E_1 \cup E_2$ is measurable.*

Proof. Let $A \subset \mathbb{R}$. Since E_2 is measurable,

$$m^*(A \cap E_1^C) = m^*(A \cap E_1^C \cap E_2) + m^*(A \cap E_1^C \cap E_2^C).$$

Then

$$\begin{aligned}
A \cap (E_1 \cup E_2) &= (A \cap E_1) \cup (A \cap E_2) \\
&= (A \cap E_1) \cup (A \cap E_2 \cap E_1^C) \\
\implies m^*(A \cap (E_1 \cup E_2)) &\leq m^*(A \cap E_1) + m^*(A \cap E_2 \cap E_1^C) \\
&= m^*(A) - m^*(A \cap E_1^C) + m^*(A \cap E_2 \cap E_1^C) \\
&= m^*(A) - m^*(A \cap E_1^C \cap E_2^C) \\
&= m^*(A) - m^*(A \cap (E_1 \cup E_2)^C) \\
\implies m^*(A \cap (E_1 \cup E_2)) + m^*(A \cap (E_1 \cup E_2)^C) &\leq m^*(A).
\end{aligned}$$

□

Theorem 7.7. *If $E_1, \ldots, E_n$ is a countable collection of measurable sets, then $\cup_{k=1}^n E_k$ is measurable.*

Proof. In the $n = 1$ case, this is clear. Suppose the claim is true for the $n - 1$ case. Let $E_1, \ldots, E_n$ be measurable. Then

$$\bigcup_{k=1}^{n} E_k = \left(\bigcup_{k=1}^{n-1} E_k \right) \cup E_n.$$

The inductive step is proven using the previous theorem. □

Definition 7.8. A nonempty collection of sets $A \subset 2^{\mathbb{R}}$ is an algebra if

1. $E \in A \implies E^C \in A$,
2. $E_1, \ldots, E_n \in A \implies \bigcup_{k=1}^n E_k \in A$.

We say an algebra A is a σ-algebra if

3. when $\{E_n\}_{n=1}^\infty \subset A$ is countable, $\bigcup_{n=1}^\infty E_n \in A$.

By De Morgan's law, for $E_1, \ldots, E_n \in A$,

$$\bigcap_{k=1}^{n} E_k = \left(\bigcup_{k=1}^{n} E_k^C \right)^C \in A.$$

Thus, if $E \in A$, $E^C \in A$, so $E \cap E^C = \emptyset \in A$. Thus, $\mathbb{R} = \emptyset^C \in A$. Similarly, if A is a σ-algebra, and $\{E_n\} \subset A$, then $\bigcup_{n=1}^\infty E_n \in A$. As is expected, we will show that $\mathcal{M} = \{\text{measurable subsets of } \mathbb{R}\} \subset 2^{\mathbb{R}}$ is a σ-algebra.

Example 7.9. The set $\{\emptyset, \mathbb{R}\}$ is a σ-algebra. The power set $2^{\mathbb{R}}$ is a σ-algebra. Let $A = \{E \subset \mathbb{R} \mid E \text{ is countable or } E^C \text{ is countable}\}$. Then A is a σ-algebra. If $\{E_n\} \subset A$, with each E_n being countable, from analysis, we proved that

$$\bigcup_{n=1}^{\infty} E_n$$

is countable, since it is a countable union of countable sets. If there is some n_0 such that $E_{n_0}^C$ is countable, then

$$\left(\bigcup_{n=1}^{\infty} E_n \right)^C = \bigcap_n E_n^C \subset E_{n_0}^C,$$

which implies that the complement of the union over $\{E_n\}$ is countable, so $\bigcup_{n=1}^\infty E_n \in A$.

Theorem 7.10. *Let*

$$\Sigma = \{A \mid A \textit{ is a } \sigma\textit{-algebra containing all open subsets of } \mathbb{R}\}.$$

Define

$$\mathcal{B} = \bigcap_{A \in \Sigma} A.$$

Then $\mathcal{B}$ is the smallest σ-algebra containing all the open subsets of $\mathbb{R}$.

We call this the Borel σ-algebra. Then this states that $\mathcal{B} \in \Sigma$, and $\forall A \in \Sigma, \mathcal{B} \subset A$.

Proof. We will partially prove this, but the unproven bits are not hard to prove. We need to verify $\mathcal{B}$ is a σ-algebra. Suppose $E \in \mathcal{B}$. Then $\forall A \in \Sigma, E \in A \implies E^C \in A \implies E^C \in \bigcap_{A \in \Sigma} A = \mathcal{B}$. □

We will continue our discussion on Lebesgue measurable subsets by proving $\mathcal{M}$ is a σ-algebra containing $\mathcal{B}$.

8 Lecture 8. Lebesgue Measurable Subsets, Measure

As a refresher, recall that E is Lebesgue measurable if $\forall A \subset \mathbb{R}$, $m^*(A) = m^*(A \cap E) + m^*(A \cap E^C)$. We also denote $\mathcal{M} = \{E \subset \mathbb{R} \mid E \text{ is Lebesgue measureble}\}$. Now, we've shown that $\mathcal{M}$ is an algebra. Recall that $\mathcal{B}$ is the Borel σ-algebra, which is the smallest σ-algebra containing all open subsets of $\mathbb{R}$. We will show that $\mathcal{M}$ is a σ-algebra.

Lemma 8.1. *Let A be an algebra, and let $\{E_n\}$ be a countable collection of elements of A. Then there exists $\{F_n\}$ a countable collection of elements of A that are disjoint such that*

$$\bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} F_n.$$

To prove that an algebra A is a σ-algebra, this lemma allows us to check that countable collections of disjoint sets of A are closed under union, since we have that countable collections of sets in A correspond to countable collections of disjoint sets.

Proof. Suppose $\{E_n\}$ is a countable collection of sets. Let

$$G_n = \bigcup_{k=1}^{n} E_k.$$

Thus, $G_1 \subset G_2 \subset \dots$, and

$$\bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} G_n.$$

Take $F_1 = G_1$, and

$$F_{n+1} = G_{n+1} \setminus G_n, \quad \forall n \geq 1.$$

Then

$$\bigcup_{k=1}^{\infty} F_k = \bigcup_{k=1}^{\infty} G_k.$$

□

Theorem 8.2. *Let $A \subset \mathbb{R}$. Let $E_1, \ldots, E_n$ be a disjoint, finite collection of measurable subsets of $\mathbb{R}$. Then*

$$m^*\left(A \cap \left[\bigcup_{k=1}^{n} E_k\right]\right) = \sum_{k=1}^{n} m^*(A \cap E_k).$$

Proof. We proceed by induction on n. In the $n = 1$ case, this is clear. Assume that our claim is true for the $n-1$ case. Let $A \subset \mathbb{R}$. Since $E_k \cap E_n = \emptyset, \forall 1, \ldots, m$,

$$A \cap \left[\bigcup_{k=1}^{n} E_k\right] \cap E_n = A \cap E_n,$$

and

$$A \cap \left[\bigcup_{k=1}^{n} E_k\right] \cap E_n^C = A \cap \left[\bigcup_{k=1}^{n-1} E_k\right].$$

Since E_n is measurable,

$$\begin{aligned} m^*(A \cap [\cup_{k=1}^{n} E_k]) &= m^*(A \cap [\cup_{k=1}^{n} E_k] \cap E_n) + m^*(A \cap [\cup_{k=1}^{n} E_k] \cap E_n^C) \\ &= m^*(A \cap E_n) + m^*(A \cap [\cup_{k=1}^{n-1} E_k]) \\ &= m^*(A \cap E_n) + \sum_{k=1}^{n-1} m^*(A \cap E_k) \\ &= \sum_{k=1}^{n} m^*(A \cap E_k). \end{aligned}$$

□

Theorem 8.3. *The set $\mathcal{M}$ is a σ-algebra.*

Proof. We've already proven $\mathcal{M}$ is an algebra, and so, by our earlier lemma, we only need to show that $\mathcal{M}$ is closed under taking countable disjoint unions. Let $\{E_n\}$ be a countable collection of disjoint measurable sets. Let $A \subset \mathbb{R}$, and let $E = \cup_{n=1}^{\infty} E_n$. We want to show that

$$m^*(A \cap E^C) + m^*(A \cap E) \le m^*(A).$$

Let $N \in \mathbb{Z}^+$. Since $\mathcal{M}$ is an algebra, $\cup_{n=1}^{N} E_n \in \mathcal{M}$. Thus,

$$\begin{aligned} m^*(A) &= m^*(A \cap \left[\cup_{n=1}^{N} E_n\right]) + m^*(A \cap \left[\cup_{n=1}^{N} E_n\right]^C) \\ &\ge m^*(A \cap \left[\cup_{n=1}^{N} E_n\right]) + m^*(A \cap E^C) \\ &= \sum_{n=1}^{N} m^*(A \cap E_n) + m^*(A \cap E^C). \end{aligned}$$

Let $N \to \infty$, then

$$\begin{aligned} m^*(A) &\ge \sum_{n=1}^{\infty} m^*(A \cap E_n) + m^*(A \cap E^C) \\ &\ge m^*(\cup_n^{\infty} A \cap E_n) + m^*(A \cap E^C) \\ &= m^*(A \cap E) + m^*(A \cap E^C). \end{aligned}$$

□

Theorem 8.4. *For all $a \in \mathbb{R}$, (a, ∞) is measurable.*

Proof. Let $A \subset \mathbb{R}$, and let $A_1 = A \cap (a, \infty)$, and $A_2 = A \cap (-\infty, a]$. We want to show that $m^*(A_1) + m^*(A_2) \leq m^*(A)$. If $m^*(A)$ is infinite, then we are done. Suppose that $m^*(A) < \infty$. Let $\varepsilon > 0$ be given, and let $\{I_n\}$ be a collection of open intervals covering A, and

$$\sum_{n=1}^{\infty} \ell(I_n) \leq m^*(A) + \varepsilon.$$

Define $J_n = I_n \cap (a, \infty)$, and $K_n = I_n \cap (-\infty, a]$. Then each J_n, K_n is an interval (possibly empty). Then

$$A_1 \subset \bigcup_{n=1}^{\infty} J_n, \quad A_2 \subset \bigcup_{n=1}^{\infty} K_n,$$

and $\ell(I_n) = \ell(J_n) + \ell(K_n)$. Then

$$\begin{aligned} m^*(A_1) + m^*(A_2) &\leq \sum_{n=1}^{\infty} m^*(J_n) + \sum_{n=1}^{\infty} m^*(K_n) \\ &= \sum_{n=1}^{\infty} \ell(J_n) + \ell(K_n) \\ &= \sum_{n=1}^{\infty} \ell(I_n) \\ &\leq m^*(A) + \varepsilon. \end{aligned}$$

□

Theorem 8.5. *Every open set of $\mathbb{R}$ is measurable, and thus $\mathcal{B} \subset \mathcal{M}$.*

Proof. For all $b \in \mathbb{R}$,

$$\begin{aligned} (-\infty, b) &= \bigcup_{n=1}^{\infty} \left(-\infty, b - \frac{1}{n}\right) \\ &= \bigcup_{n=1}^{\infty} \left(b - \frac{1}{n}, \infty\right)^C \in \mathcal{M}. \end{aligned}$$

Thus, $\forall a, b \in \mathbb{R}$, $(a, b) = (-\infty, b) \cap (a, \infty)$ is measurable, since σ-algebras are closed under countable intersections. Since every open subset of $\mathbb{R}$ is a countable union of disjoint open intervals, our claim is proven. □

8.1 Lebesgue Measure

Definition 8.6. If $E \in \mathcal{M}$, the Lebesgue measure of E is $m(E) = m^*(E)$.

Theorem 8.7. *If $A, B \in \mathcal{M}$ and $A \subset B$, then $m(A) \leq m(B)$.*

Theorem 8.8. *If $I \subset \mathbb{R}$ is an interval, then $I \in \mathcal{M}$, and $m(I) = \ell(I)$.*

Proof. Consider that $[a, b] = (b, \infty)^C \cap (-\infty, a)^C \in \mathcal{M}$. □

Theorem 8.9. *Suppose $\{E_n\}$ is a countable collection of disjoint measurable sets. Then*

$$m\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} m(E_n).$$

Proof. We always have

$$m\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} m(E_n),$$

so we want to show the reverse inequality. Let $N \in \mathbb{Z}^+$. Then consider

$$\begin{aligned}
m\left(\bigcup_{n=1}^{N} E_n\right) &= m^*\left(\mathbb{R} \cap \left[\bigcup_{k=1}^{N} E_n\right]\right) \\
&= \sum_{n=1}^{N} m^*(\mathbb{R} \cap E_n) \\
&= \sum_{n=1}^{N} m(E_n) \\
\implies \sum_{n=1}^{N} m(E_n) &= m\left(\bigcup_{n=1}^{N} E_n\right) \\
&\leq m\left(\bigcup_{n=1}^{\infty} E_n\right) \\
\implies \sum_{n=1}^{\infty} m(E_n) &\leq m\left(\bigcup_{n=1}^{\infty} E_n\right). && (n \to \infty)
\end{aligned}$$

□

Theorem 8.10. *If $E \in \mathcal{M}$, and $x \in \mathbb{R}$, then $E+x = \{y+x \mid y \in E\}$ is measurable, and $m(E) = m(E+x)$.*

Theorem 8.11 (Continuity of Measure)**.** *Suppose $\{E_k\}$ is a collection of measurable sets s.t. $E_1 \subset E_2 \subset E_3 \subset \ldots$. Then*

$$m\left(\bigcup_{k=1}^{\infty} E_k\right) = \lim_{n\to\infty} m\left(\bigcup_{k=1}^{n} E_k\right) = \lim_{n\to\infty} m(E_n).$$

Proof. Let $F_1 = E_1$, $F_{k+1} = E_{k+1} \setminus E_k = E_{k+1} \cap E_k^C, \forall k \geq 1$. Note that each F_k is measurable. Then $\{F_k\}$ is a disjoint union of measurable sets, and $\forall n \in \mathbb{Z}^+$,

$$\bigcup_{k=1}^{n} F_k = E_n.$$

Also,

$$\bigcup_{k=1}^{\infty} F_k = \bigcup_{k=1}^{\infty} E_k.$$

Then

$$\begin{aligned}
m\left(\bigcup_{k=1}^{\infty} E_k\right) &= m\left(\bigcup_{k=1}^{\infty} F_k\right) \\
&= \sum_{k=1}^{\infty} m(F_k) \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} m(F_n) \\
&= \lim_{n\to\infty} m\left(\bigcup_{k=1}^{n} F_k\right) \\
&= \lim_{n\to\infty} m(E_n).
\end{aligned}$$

□

A similar proof is used to prove the analog result, when $\{E_n\}$ is a decreasing sequence of measurable sets.

9 Lecture 9. Lebesgue Measurable Functions

9.1 Measurable Functions

Riemann integration involves partitioning the domain of a function, evaluating the function at some nicely chosen point at each part, and evaluating a Riemann sum. When Lebesgue invented his theory of integration, he was interested in chopping up the range of the function being integrated, and studying the inverse images of certain parts of the range. We then generalize by instead of studying parts which might look like intervals, studying Lebesgue measurable sets of the range. To get to this point, we should naturally like to study functions whose inverse images of certain sets are measurable.

Definition 9.1. We've been working implicitly in the extended real-number line, so we should extend this idea to functions. Write $[-\infty, \infty] = \mathbb{R} \cup \{\pm\infty\}$. Sums are defined as $x + \pm\infty := \pm\infty$, $\forall x \in \mathbb{R}$. We still avoid indeterminate forms like $\infty - \infty$. Products are defined as $0 \cdot (\pm\infty) := 0$, and $x \cdot (\pm\infty) := \pm\infty, \forall x \in \mathbb{R}^*$. Recall that we say $\{a_n\} \subset \mathbb{R}$ converges to ∞ if $\forall M > 0$, $\exists N \in \mathbb{Z}^+$ such that $\forall n \geq N$, $a_n > M$. Similarly for when $a_n \to -\infty$.

Let's define what it means for a function to be Lebesgue measurable.

Definition 9.2. Let $E \subset \mathbb{R}$ be measurable, and $f : E \to [-\infty, \infty]$. We say f is measurable if $\forall \alpha \in \mathbb{R}$,

$$f^{-1}((\alpha, \infty]) \in \mathcal{M}.$$

Theorem 9.3. *Let $E \subset \mathbb{R}$ be measurable, and $f : E \to [-\infty, \infty]$. TFAE:*

1. $\forall \alpha \in \mathbb{R}, f^{-1}((\alpha, \infty]) \in \mathcal{M}$,
2. $\forall \alpha \in \mathbb{R}, f^{-1}([\alpha, \infty]) \in \mathcal{M}$,
3. $\forall \alpha \in \mathbb{R}, f^{-1}([-\infty, \alpha)) \in \mathcal{M}$.

4. $\forall \alpha \in \mathbb{R}, f^{-1}([-\infty, \alpha]) \in \mathcal{M}$.

Proof. Assume 1. Then $\forall \alpha \in \mathbb{R}$,

$$[\alpha, \infty) = \bigcap_{n=1}^{\infty} \left(\alpha - \frac{1}{n}, \infty\right]$$
$$\implies f^{-1}([\alpha, \infty]) = \bigcap_{n=1}^{\infty} f^{-1}\left(\left(\alpha - \frac{1}{n}, \infty\right)\right) \in \mathcal{M}.$$

Assume 2. Then $\forall \alpha \in \mathbb{R}$,

$$(\alpha, \infty] = \bigcup_{n=1}^{\infty} \left[\alpha + \frac{1}{n}, \infty\right]$$
$$\implies f^{-1}((\alpha, \infty]) = \bigcup_{n=1}^{\infty} f^{-1}\left(\left[\alpha + \frac{1}{n}, \infty\right]\right) \in \mathcal{M}.$$

We have 2 $\iff$ 3 since $\forall \alpha \in \mathbb{R}$,

$$[-\infty, \alpha) = ([\alpha, \infty])^C,$$

and 1 $\iff$ 4 since $\forall \alpha \in \mathbb{R}$,

$$[-\infty, \alpha] = ((\alpha, \infty])^C.$$

□

Theorem 9.4. *If E is measurable, and $f : E \to \mathbb{R}$ is a measurable function, then $\forall F \in \mathcal{B}$, the Borel σ-algebra, $f^{-1}(F)$ is measurable.*

Proof. If f is measurable, then $\forall a, b \in \mathbb{R}, a < b$,

$$f^{-1}((a, b)) = f^{-1}([-\infty, b) \cap (a, \infty]) = f^{-1}([-\infty, b)) \cap f^{-1}((a, \infty]) \in \mathcal{M}.$$

Thus, $\forall a < b$, $f^{-1}((a, b))$ is measurable. Thus, $f^{-1}(U)$ is measurable, for all $U \subset \mathbb{R}$ being open. Since $A = \{F \subset \mathbb{R} \mid f^{-1}(F) \text{ is measurable}\}$ is a σ-algebra containing all open sets (homework), $\mathcal{B} \subset A$. □

Theorem 9.5. *If $f : E \to \mathbb{R}$ is measurable, then $f^{-1}(\{\infty\})$ and $f^{-1}(\{-\infty\})$ are measurable.*

Proof. We have that $f^{-1}(\{\infty\} = \bigcap_{n=1}^{\infty} f^{-1}((n, \infty]) \in \mathcal{M}$. Similarly $f^{-1}(\{-\infty\}) = \bigcap_{n=1}^{\infty} f^{-1}([-\infty, n)) \in \mathcal{M}$. □

What are the simplest types of measurable functions?

Example 9.6. If $f : \mathbb{R} \to \mathbb{R}$ is continuous, then f is measurable. Consider that $\forall \alpha \in \mathbb{R}$,

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

since it is open.

Example 9.7. Let $E, F \subset \mathbb{R}$ be measurable, and define

$$\chi_F(x) = \begin{cases} 1 & x \in F \\ 0 & x \notin F. \end{cases}$$

Then $\chi_F : E \to \mathbb{R}$ is measurable. If $\alpha \in \mathbb{R}$,

$$\chi_F^{-1}((\alpha, \infty]) = \begin{cases} \emptyset & \alpha \geq 1 \\ E \cap F & 0 \leq \alpha < 1 \\ E & \alpha < 0 \end{cases}.$$

Theorem 9.8. *If $E \subset \mathbb{R}$ is measurable, and $f, g : E \to \mathbb{R}$ are measurable, with $c \in \mathbb{R}$. Then $cf, f+g, fg : E \to \mathbb{R}$ are measurable.*

Proof. If $c = 0$, then $cf = 0$, which is continuous, hence, measurable. Suppose $c > 0$, and let $\alpha \in \mathbb{R}$. Then

$$\begin{aligned} cf(x) > \alpha &\implies f(x) > \frac{\alpha}{c} \\ &\implies (cf)^{-1}((\alpha, \infty]) = f^{-1}([\alpha/c, \infty]) \in \mathcal{M}. \end{aligned}$$

A similar calculation follows for $c < 0$. Suppose $\alpha \in \mathbb{R}$. Then

$$\begin{aligned} f(x) + g(x) > \alpha &\implies f(x) > \alpha - g(x) \\ &\implies \exists r \in \mathbb{Q}, f(x) > r > \alpha - g(x) \\ &\implies \exists r \in \mathbb{Q}, x \in f^{-1}((r, \infty]) \cap g^{-1}((\alpha - r, \infty)) \in \mathcal{M}. \end{aligned}$$

I.e.

$$(f+g)^{-1}((\alpha, \infty]) = \bigcup_{r \in \mathbb{Q}} [f^{-1}((r, \infty]) \cap g^{-1}((\alpha - e, \infty])] \in \mathcal{M},$$

since $\mathbb{Q}$ is countable. Now, we claim that f^2 is measurable. Let $\alpha \in \mathbb{R}$. If $\alpha < 0$, then $(f^2)^{-1}((\alpha, \infty)) = E \in \mathcal{M}$. Assume that $\alpha \geq 0$, then $f^2(x) > \alpha \iff f(x) > \sqrt{\alpha} \vee f(x) < -\sqrt{\alpha}$. Thus,

$$(f^2)^{-1}((\alpha, \infty]) = f^{-1}((\sqrt{\alpha}, \infty]) \cup f^{-1}([-\infty, -\sqrt{a})) \in \mathcal{M}.$$

Then

$$fg = \frac{1}{4}[(f+g)^2 - (f-g)^2] \in \mathcal{M}.$$

□

If one of $f, g : E \to [-\infty, \infty]$, that is fine, and the proofs presented are fine. If both $f, g : E \to [-\infty, \infty]$, our product rule still holds, but we potentially are undefined in the sum case, where we might encounter $\infty - \infty$ expressions.

Theorem 9.9. *If $E \subset \mathbb{R}$ is measurable, and $f_n : E \to [-\infty, \infty]$ is measurable, then*

1. $g_1(x) = \sup_{n \in \mathbb{Z}^+} f_n(x)$,
2. $g_2(x) = \inf_{n \in \mathbb{Z}^+} f_n(x)$,

3. $g_3(x) = \limsup_{n\to\infty} f_n(x) = \inf_{n\in\mathbb{Z}^+} \sup_{k\geq n} f_k(x)$,

4. $g_4(x) = \liminf_{n\to\infty} f_n(x) = \sup_{n\in\mathbb{Z}^+} \inf_{k\geq n} f_k(x)$,

are all measurable functions.

Proof. Consider that

$$\begin{aligned} x \in g_1^{-1}((\alpha,\infty)) &\iff \sup_{n\in\mathbb{Z}^+} f_n(x) > \alpha \\ &\iff \exists n \in \mathbb{Z}^+, x \in f_n^{-1}((\alpha,\infty]) \\ \implies g_1^{-1}((\alpha,\infty]) &= \bigcup_{n\in\mathbb{Z}^+} f_n^{-1}((\alpha,\infty]) \in \mathcal{M}. \end{aligned}$$

Similarly, if $\alpha \in \mathbb{R}$,

$$g_2^{-1}([\alpha,\infty]) = \bigcap_{n\in\mathbb{Z}^+} f_n^{-1}([\alpha,\infty]) \in \mathcal{M}.$$

That g_3, g_4 are measurable follows from the fact that g_1, g_2 are measurable. □

Corollary 9.9.1. *If $E \subset \mathbb{R}$ is measurable, and $f_n : E \to [-\infty,\infty]$ is measurable, $\forall n \in \mathbb{Z}^+$, and*

$$\lim_{n\to\infty} f_n(x) = f(x), \quad \forall x \in E,$$

then f is measurable.

Proof. We have that $\forall x \in E, f(x) = \limsup_{n\in\mathbb{Z}^+} f_n(x) = \liminf_{n\in\mathbb{Z}^+} f_n(x)$. □

Note. If $f_n : [a,b] \to \mathbb{R}$ is Riemann integrable, $\forall n \in \mathbb{Z}^+$, and $f_n \to f$ pointwise, then f need not be Riemann integrable. This convergence needs to be uniform for this to hold. To briefly revisit, consider $\mathbb{Q} \cap [0,1] = \{r_1, r_2, \ldots\}$. If we set

$$f_n(x) = \begin{cases} 1 & x \in \{r_1,\ldots,r_n\} \\ 0 & \text{otherwise ,} \end{cases}$$

then f_n is Riemann integrable, since it is piecewise continuous. However, $f_n \to \chi_{\mathbb{Q}\cap[0,1]}$, which is not Riemann integrable, since it is discontinuous everywhere.

Definition 9.10. Let's introduce some terminology. A statement $P(x)$ holds almost everywhere on E if

$$m(\{x \in E \mid P(x) \text{ does not hold}\}) = 0.$$

Theorem 9.11. *If $f, g : E \to [-\infty,\infty]$, f is measurable, and $f = g$ almost everywhere on E, then g is measurable.*

Proof. Let $N = \{x \in E \mid f(x) \neq g(x)\}$. Then $m(N) = 0$, and N is measurable. If $\alpha \in \mathbb{R}$, then

$$N_\alpha := \{x \in N \mid g(x) > \alpha\} \subset N$$

so $m^*(N_\alpha) = 0$, and therefore, N_α is measurable. Then

$$g^{-1}((\alpha,\infty]) = f^{-1}((\alpha,\infty] \cap N^C) \cup N_\alpha \in \mathcal{M}.$$

□

Next lecture, we'll extend this notion of being measurable to the complex numbers, and then we'll introduce this new concept of simple functions.

10 Lecture 10. Simple Functions

Definition 10.1. Let $E \subset \mathbb{R}$ be measurable. We say $f : E \to \mathbb{C}$ is measurable if $\mathrm{Re}(f) : E \to \mathbb{R}$ and $\mathrm{Im}(f) : E \to \mathbb{R}$ are both measurable.

Proposition 10.2. *If $f, g : E \to \mathbb{C}$ are measurable and $\alpha \in \mathbb{C}$, then αf, $f+g$, fg, $\overline{f}$, and $|f|$ are measurable.*

Theorem 10.3. *If $f_n : E \to \mathbb{C}$ is measurable $\forall n \in \mathbb{Z}^+$, and*

$$\lim_{n\to\infty} f_n(x) = f(x), \quad \forall x \in E,$$

then f is measurable.

Proof. Consider that $\forall x \in E$,

$$\lim_{n\to\infty} f_n(x) = f(x) \iff \lim_{n\to\infty} \mathrm{Re}(f_n(x)) = \mathrm{Re}(f(x)) \wedge \lim_{n\to\infty} \mathrm{Im}(f_n(x)) = \mathrm{Im}(f(x)).$$

□

Let's talk about a nice class of functions.

Definition 10.4. If $E \subset \mathbb{R}$ is measurable, a measurable function $\varphi : E \to \mathbb{C}$ is a simple function if

$$\varphi(E) = \{a_1, a_2, \ldots, a_n\}, \quad a_i \in \mathbb{C}, i \in [n].$$

That is, the range of φ is finite.

Note. If φ is a simple function, with $\varphi(E) = \{a_1, \ldots, a_n\}$, then $\forall i \in [n]$, set $A_i := \varphi^{-1}(\{a_i\})$, and note that A_i is measurable. (Since singleton sets are closed, and hence, belong to $\mathcal{B}$, and measurable functions inverse images of Borel sets are measurable.) Also,

- $\forall i \neq j, A_i \cap A_j = \emptyset$,
- $\cup_{i=1}^n A_i = E$,
- $\forall x \in E$,

$$\varphi(x) = \sum_{i=1}^n a_i \chi_{A_i}(x).$$

Theorem 10.5. *Linear combinations and products of simple functions are simple.*

Theorem 10.6. *If $f : E \to [0, \infty]$ is measurable, then there exists a sequence of simple functions $\{\varphi_n\}$ such that*

1. *$\forall x \in E$,*
$$0 \le \varphi_1(x) \le \varphi_2(x) \le \varphi_3(x) \le \ldots \le f(x),$$
2. *$\forall x \in E$,*
$$\lim_{n\to\infty} \varphi_n(x) = f(x),$$
3. *$\forall B > 0$, $\varphi_n \to f$ uniformly on*
$$\{x \in E \mid f(x) \le B\}.$$

Proof. The idea is to build functions φ_n to have better and better "resolution" (2^{-n}) and larger and larger range (2^n). Essentially, φ_0 will only be able to tell whether the function is at least 1 (we'll only let it take on the values 0 and 1, being 1 is $f \geq 1$, and 0 otherwise), φ_1 will be able to tell the values of functions up to 2 (resolving at intervals of $1/2$, so that it can take on the values 0, $1/2$, 1, 3/2,2), and so on. We claim this sequence of approximations satisfies the three conditions we want above. For $n = 0, 1, 2, \ldots$, and for $0 \leq k \leq 2^{2n} - 1$, define

$$\begin{aligned} E_n^k &= \{x \in E \mid k2^{-n} < f(x) < (k+1)2^{-n}\} \\ &= f^{-1}((k2^{-n}, (k+1)2^{-n}]). \end{aligned}$$

Set

$$F_n = f^{-1}((2^n, \infty]).$$

Set

$$\varphi_n = \left(\sum_{k=0}^{2^{2n}-1} (k2^{-n}) \chi_{E_n^k} \right) + 2^n \chi_{F_n}.$$

For example

$$\varphi_1 = 0 \cdot \chi_{f^{-1}((0,1/2])} + \frac{1}{2} \chi_{f^{-1}((1/2,1])} + 1 \cdot \chi_{f^{-1}((1,3/2])} + \frac{3}{2} \chi_{f^{-1}((3/2,2])} + 2\chi_{f^{-1}((2,\infty])}.$$

By construction, it is true that φ_n takes on finitely many values for each n, so φ_n is always a simple function. Also, $0 \leq \varphi_n \leq f$. (For example, if $f(x) = 1.7$, at some point x, then we fall within the $(3/2, 2]$ range, and then φ_1 takes on the lower bound of that range $3/2$.) More rigorously, if $x \in E_n^k$, then

$$k2^{-n} < f(x) \leq (k+1)2^{-n} \Longrightarrow \varphi_n(x) = k2^{-n} < f(x),$$

and otherwise, $x \in F_n$, which means $f(x) > 2^n = \varphi_n(x)$. All that remains to for proving part 1 is to show that the φ_n's are pointwise increasing, notice that if $x \in E_n^k$, then

$$k2^{-n} < f(x) < (k+1)2^{-n} \Longrightarrow (2k)2^{-(n+1)} < f(x) \leq (2k+2)2^{-(n+1)},$$

which implies that $x \in E_{n+1}^{2k} \cup E_{n+1}^{2k+1}$. And we can check in both cases that $\varphi_{n+1}(x)$ is larger than $\varphi_n(x)$. If $x \in E_{n+1}^{2k}$, then

$$\varphi_n(x) = k2^{-n} = (2k)2^{-(n+1)} = \varphi_{n+1}(x),$$

and otherwise, $x \in E_{n+1}^{2k+1}$, which means that

$$\varphi_n(x) = k2^{-n} = (2k)2^{-n+1} < (2k+1)2^{-(n+1)} = \varphi_{n+1}(x).$$

Finally, if $x \in F_n$, then $\varphi_n(x) \leq \varphi_{n+1}(x)$, by a similar argument. So we've shown that $\varphi_n(x) \leq \varphi_{n+1}(x)$ on each of the sets F_n and E_n^k, which partition E, and thus, part 1 is proven.

We can now prove parts 2, 3 by the following. We claim that for all $x \in \{y \in E \mid f(y) \leq 2^n\}$,

$$0 \leq f(x) - \varphi_n(x) \leq 2^{-n}.$$

Once we show this claim, we can show part 2 because for any x, either $f(x) = \infty$, or $f(x)$ is in the sets $\{y \in E \mid f(y) \leq 2^n\}$ for n large enough, so then for sufficiently large n, we have

$|f(x) - \varphi_n(x)| \leq 2^{-n}$, which is enough for pointwise convergence. Part 3 follows because for any fixed B, we can pick an N so that $\{x \in E \mid f(x) \leq B\}$ is contained in $\{x \in E \mid f(x) \leq 2^N\}$, and then the bound in the claim also shows uniform convergence.

In order to prove our claim, remember that φ_n cuts up our range into intervals of resolution 2^{-n}. Since

$$\{y \in E \mid f(x) \leq 2^n\} = \bigcup_{k=0}^{2^{2n-1}} E_n^k,$$

we can just check the claim holds on each individual E_n^k. If $x \in E_n^k$, then

$$\varphi_n(x) = k2^{-n} \leq f(x) \leq (k+1)2^{-n} \implies f(x) - \varphi_n(x) \leq (k+1)2^{-n} - k2^{-n} = 2^{-n},$$

as desired, completing the proof. □

This theorem carries over without difficulty to complex functions as follows.

Definition 10.7. If $f : E \to [-\infty, \infty]$, we define $f^+(x) = \max(f(x), 0)$ (the positive part), and $f^-(x) = \max(-f(x), 0)$ (the negative part). Note that f^+, f^- are nonnegative.

Theorem 10.8. *Let $f : E \to [-\infty, \infty]$. Then $f = f^+ - f^-$, and $|f| = f^+ + f^-$.*

Theorem 10.9. *Let $E \subset \mathbb{R}$ be measurable, and $f : E \to \mathbb{C}$ be measurable. Then there exists a sequence of simple functions $\{\varphi_n\}$ such that*

1. *$\forall x \in E$,*
$$0 \leq |\varphi_0(x)| \leq |\varphi_1(x)| \leq |\varphi_2(x)| \leq \ldots \leq |f(x)|,$$
2. *$\forall x \in E$,*
$$\lim_{n\to\infty} \varphi_n(x) = f(x),$$
3. *$\forall B \geq 0$, $\varphi_n \to f$ uniformly on*
$$\{x \in E \mid |f(x)| \leq B\}.$$

Proof. Apply Theorem 10.6 to the real/imaginary positive/negative parts of f. The linear combinations of the simple functions that arise from each of those parts will give us the desired approximation for f. □

10.1 Lebesgue Integral of Simple Nonnegative Functions

Definition 10.10. If $E \subset \mathbb{R}$ is measurable, we define

$$L^+(E) = \{f : E \to [0, \infty] \mid f \text{ is measurable}\}.$$

Definition 10.11. Let $\varphi \in L^+(E)$ be a simple function, such that

$$\varphi = \sum_{j=1}^{n} a_j \chi_{A_j},$$

where $\forall j \in [n], A_j \subset E, \forall i \neq j, A_i \cap A_j = \emptyset$, and $\cup_{j=1}^n A_j = E$. The Lebesgue integral of φ

$$\int_E \varphi := \sum_{j=1}^{n} a_j m(A_j) \in [0, \infty].$$

We may also write $\int_E \varphi(x)\, dx$.

Theorem 10.12. *Let $\varphi, \psi \in L^+(E)$ be a simple function.*

1. *If $c \geq 0$, then $\int_E c\varphi = c\int_E \varphi$.*
2. *Linearity: $\int_E(\varphi + \psi) = \int_E \varphi + \int_E \psi$.*
3. *If $\varphi \leq \psi$ $(\forall x \in E, \varphi(x) \leq \psi(x))$, then*
$$\int_E \varphi \leq \int_E \psi.$$
4. *If $F \subset E$ is measurable, then*
$$\int_F \varphi = \int_E \chi_F \varphi \leq \int_E \varphi.$$

Proof. Consider

$$\begin{aligned} c\varphi &= \sum_{j=1}^n (ca_j)\chi_{A_j} \\ \implies \int_E c\varphi &= \sum_{j=1}^n ca_j m(A_j) \\ &= c\sum_{j=1}^n a_j m(A_j) = c\int_E \varphi. \end{aligned}$$

For part 2, write

$$\varphi = \sum_{j=1}^n a_j \chi_{A_j}, \quad \psi = \sum_{k=1}^m b_k \chi_{B_k}.$$

Then

$$\begin{aligned} E &= \bigcup_{j=1}^n A_j = \bigcup_{k=1}^n B_k \\ \implies \forall j \in [n], A_j &= \bigcup_{k=1}^m A_j \cap B_k, \\ \forall k \in [m], B_k &= \bigcup_{j=1}^n B_k \cap A_j, \end{aligned}$$

and note that these unions are disjoint. Then by additivity of m,

$$\begin{aligned} \int_E \varphi + \int_E \psi &= \sum_{j=1}^n a_j m(A_j) + \sum_{k=1}^m b_k m(B_k) \\ &= \sum_{j,k} a_j m(A_j \cap B_k) + \sum_{k,j} b_k m(A_j \cap B_k) \\ &= \sum_{j,k} (a_j + b_k) m(A_j \cap B_k). \end{aligned}$$

Since $\varphi + \psi = \sum_{j,k}(a_j + b_k)\chi_{A_j \cap B_k}$,

$$\int_E (\varphi + \psi) = \sum_{j,k}(a_j + b_k)m(A_j \cap B_k) = \int_E \varphi + \int_E \psi.$$

For part 3,

$$\forall x \in E, \varphi(x) \leq \psi(x) \Longrightarrow a_j \leq b_k \text{ whenever } A_j \cap B_k \neq \emptyset.$$

Thus,

$$\begin{aligned} \int_E \varphi &= \sum_{j=1}^{n} a_j m(A_j) \\ &= \sum_{j,k} a_j m(A_j \cap B_k) \\ &\leq \sum_{j,k} b_k m(A_j \cap B_k) \\ &= \int_E \psi. \end{aligned}$$

Part 4 is a simple exercise left to the motivated reader. □

We've now defined the "area under the curve" for Lebesgue integrals, and this is an indication that Riemann integrable functions will indeed be Lebesgue integrable (since step functions are indeed Riemann integrable). Next lecture, we will define the integral of nonnegative measurable functions, and we will prove two important convergence theorems.

11 Lecture 11. Lebesgue Integral of Nonnegative Functions, Convergence Theorems

Definition 11.1. If $f \in L^+(E)$, we define

$$\int_E f = \sup\left\{\int_E \varphi \mid \varphi \in L^+(E), \varphi \text{ is simple}, \varphi \leq f\right\}.$$

Theorem 11.2. *If $E \subset \mathbb{R}$ is measurable with $m(E) = 0$, then $\forall f \in L^+(E)$, $\int_E f = 0$.*

Proof. Let $\varphi \in L^+(E)$ be simple,

$$\varphi = \sum_{j=1}^{n} a_j \chi_{A_j},$$

with $\varphi \leq f$. Then $A_j \subset E$, $m(A_j) = 0$, so

$$\int_E \varphi = \sum_{j=1}^{n} a_j m(A_j) = 0 \Longrightarrow \int_E f = 0.$$

□

Theorem 11.3. *The following hold.*

1. *If $\varphi \in L^+(E)$ is simple, then the two definitions of $\int_E \varphi$ agree.*
2. *If $f, g \in L^+(E)$, $c \in [0, \infty)$, and $f \leq g$ on E, then $\int_E cf \leq c \int_E f$, and $\int_E f \leq \int_E g$.*
3. *If $f \in L^+(E)$ and $F \subset E$ is measurable, then $\int_F f = \int_E f\chi_F \leq \int_E f$.*

Theorem 11.4. *If $f, g \in L^+(E)$, and $f \leq g$ almost everywhere on E, then*

$$\int_E f \leq \int_E g.$$

Proof. Let $F = \{x \in E \mid f(x) \leq g(x)\}$. Then F is measurable (left to reader), and $m(F^C) = 0$. Then

$$\int_E f = \int_{F \cup F^C} f = \int_F f + \int_{F^C} f.$$

(We haven't shown why this is true, but it is a generalization of the simple function case.) Then

$$\int_E f = \int_F f \leq \int_F g = \int_F g + \int_{F^C} g = \int_E g.$$

□

11.1 Monotone Convergence Theorem

Theorem 11.5 (Monotone Convergence Theorem)**.** *If $\{f_n\}$ is a sequence in $L^+(E)$ such that*

$$f_1 \leq f_2 \leq f_3 \leq \dots$$

on E, and $f_n \to f$ pointwise on E, then

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

Proof. If $f_1 \leq f_2 \leq \dots$,

$$\int_E f_1 \leq \int_E f_2 \leq \dots.$$

Therefore, $\lim_{n\to\infty} \int_E f_n$ exists in $[0, \infty]$. Since $f_1 \leq f_2 \leq \dots$, and $\lim_{n\to\infty} f_n(x) = f(x)$, $\forall x \in E$, $f_1 \leq f_2 \leq \dots \leq f$. Therefore, $\forall n \in \mathbb{Z}^+$,

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

Now we show that

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

Let $\varphi \in L^+(E)$ be simple, with

$$\varphi = \sum_{j=1}^{m} a_j \chi_{A_j},$$

with $\varphi \leq f$. Let $\varepsilon \in (0, 1)$, and

$$E_n = \{x \in E \mid f_n(x) \geq (1 - \varepsilon)\varphi(x)\}.$$

Note that $\forall x \in E, (1-\varepsilon)\varphi(x) < f(x)$. Then since for all $x \in E$,

$$\lim_{n\to\infty} f_n(x) = f(x) \Longrightarrow \bigcup_{n=1}^{\infty} E_n = E.$$

Since $f_1 \leq f_2 \leq f_3 \leq \dots$, we have $E_1 \subset E_2 \subset E_3 \subset \dots$. Consider that

$$\begin{aligned}\int_E f_n \geq \int_{E_n} f_n \geq \int_{E_n} (1-\varepsilon)\varphi \\ &= (1-\varepsilon)\int_{E_n} \varphi \\ &= (1-\varepsilon)\sum_{j=1}^{m} a_j m(A_j \cap E_n).\end{aligned}$$

Since $E_1 \cap A_j \subset E_2 \cap A_j \subset \dots$, and

$$\bigcup_{n=1}^{\infty}(E_n \cap A_j) = A_j \Longrightarrow \forall j, \lim_{n\to\infty} m(A_j \cap E_n) = m\left(\bigcup_{n=1}^{\infty} E_n \cap A_j\right) = m(A_j).$$

Thus,

$$\begin{aligned}\lim_{n\to\infty}\int_E f_n &\geq \lim_{n\to\infty}(1-\varepsilon)\sum_{j=1}^{m} a_j m(A_j \cap E_n) \\ &= (1-\varepsilon)\sum_{j=1}^{m} a_j m(A_j) \\ &= (1-\varepsilon)\int_E \varphi.\end{aligned}$$

Thus, $\forall \varepsilon \in (0,1)$,

$$(1-\varepsilon)\int_E \varphi \leq \lim_{n\to\infty}\int_E f_n \Longrightarrow \int_E \varphi \leq \lim_{n\to\infty}\int_E f_n.$$

Since φ is an arbitrary simple function,

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

□

Theorem 11.6. *If $f \in L^+(E)$, and $\{\varphi_n\}$ is a sequence of simple functions such that*

$$0 \leq \varphi_1 \leq \varphi_2 \leq \dots \leq f,$$

and $\varphi_n \to f$ pointwise, then

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

Theorem 11.7. *If $f, g \in L^+(E)$, then*

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

Proof. Let $\{\varphi_n\}$, $\{\psi_n\}$ be two sequences of simple functions such that

$$0 \le \varphi_1 \le \ldots \le f,$$

with $\varphi_n \to f$ pointwise, and

$$0 \le \psi_1 \le \ldots \le g,$$

with $\psi_n \to g$ pointwise. Then

$$0 \le \varphi_1 + \psi_1 \le \ldots \le f + g,$$

with $\varphi_n + \psi_n \to f + g$ pointwise. Then

$$\begin{aligned} \int_E (f+g) &= \lim_{n\to\infty} \int_E \varphi_n + \psi_n \\ &= \lim_{n\to\infty} \left[\int_E \varphi_n + \int_E \psi_n \right] \\ &= \int_E f + \int_E g. \end{aligned}$$

□

Theorem 11.8. *If* $\{f_n\} \subset L^+(E)$*, then*

$$\int_E \sum_{n=1}^{\infty} f_n = \sum_{n=1}^{\infty} \int_E f_n$$

Proof. By an induction argument, and the previous theorem on linearity of the Lebesgue integral, we have $\forall N \in \mathbb{Z}^+$,

$$\int_E \sum_{n=1}^{N} f_n = \sum_{n=1}^{N} \int_E f_n.$$

Since

$$\sum_{n=1}^{1} f_n \le \sum_{n=1}^{2} f_n \le \ldots,$$

and $\sum_{n=1}^{n} f_n \to \sum_{n=1}^{\infty} f_n$ pointwise as N increases to infinity, by the monotone convergence theorem,

$$\begin{aligned} \int_E \sum_{n=1}^{\infty} f_n &= \lim_{N\to\infty} \int_E \sum_{n=1}^{N} f_n \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} \int_E f_n \\ &= \sum_{n=1}^{\infty} \int_E f_n. \end{aligned}$$

□

Theorem 11.9. *If $f \in L^+(E)$, then*

$$\int_E f = 0 \Longleftrightarrow f = 0 \text{ almost everywhere on } E.$$

Proof. If $f = 0$ almost everywhere, then

$$0 \leq \int_E f \leq \int_E 0 = 0.$$

Assume that $\int_E f = 0$. Let $F_n = \{x \in E \mid f(x) > 1/n\}$, and $F = \{x \in E \mid f(x) > 0\}$. Then

$$\bigcup_{n=1}^{\infty} F_n = F,$$

and $F_1 \subset F_2 \subset F_3 \subset \dots$. Then $\forall n \in \mathbb{Z}^+$,

$$0 \leq \frac{1}{n} m(F_n) = \int_{F_n} \frac{1}{n} \chi_{F_n} \leq \int_{F_n} f \leq \int_E f = 0.$$

Therefore, $\forall n \in \mathbb{Z}^+$, $m(F_n) = 0$, so

$$m(F) = m\left(\bigcup_{n=1}^{\infty} F_n\right) = \lim_{n\to\infty} m(F_n) = 0.$$

Thus, $f = 0$ almost everywhere. □

Theorem 11.10. *If $\{f_n\}$ is a sequence in $L^+(E)$ such that for almost every $x \in E$,*

$$(*) \begin{cases} f_1(x) \leq f_2(x) \leq \dots, \\ \lim_{n\to\infty} f_n(x) = f(x). \end{cases}$$

Then

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

Proof. Let $F = \{x \in E \mid (*) \text{ holds }\}$. Then $m(F^C) = m(E \setminus F) = 0$. Then $f - \chi_F f = 0$ almost everywhere, and $f_n - \chi_F f_n = 0$ almost everywhere, for all $n \in \mathbb{Z}^+$. By the monotone convergence theorem,

$$\begin{aligned} \int_E f &= \int_E f \chi_F \\ &= \int_F f \\ &= \lim_{n\to\infty} \int_F f_n \\ &= \lim_{n\to\infty} \int_E f_n. \end{aligned}$$

□

11.2 Fatou's Lemma

Theorem 11.11. *If $\{f_n\}$ is a sequence in $L^+(E)$, then*

$$\int_E \liminf_{n\to\infty} f_n(x) \leq \liminf_{n\to\infty} \int_E f_n.$$

Proof. We have

$$\begin{aligned}\liminf_{n\to\infty} f_n(x) &= \sup_{n\geq 1} \inf_{k\geq n} f_k(x) \\ &= \lim_{n\to\infty} \inf_{k\geq n} f_k(x).\end{aligned}$$

and

$$\inf_{k\geq 1} f_k(x) \leq \inf_{k\geq 2} f_k(x) \leq \dots,$$

by the monotone convergence theorem,

$$\int_E \liminf_{n\to\infty} f_n = \lim_{n\to\infty} \int_E \inf_{k\geq n} f_k.$$

For all $j \geq n$, $\forall x \in E$,

$$\inf_{k\geq n} f_k(x) \leq f_j(x).$$

Therefore, $\forall j \geq n$,

$$\int_E \inf_{k\geq n} f_k \leq \int_E f_j \Longrightarrow \int_E \inf_{k\geq n} f_k \leq \inf_{j\geq n} \int_E f_j.$$

Thus

$$\int_E \liminf_{n\to\infty} f_n = \lim_{n\to\infty} \int_E \inf_{k\geq n} f_k \leq \lim_{n\to\infty} \inf_{j\geq n} \int_E f_j = \liminf_{n\to\infty} \int_E f_n.$$

□

Theorem 11.12. *If $f \in L^+(E)$, and*

$$\int_E f < \infty,$$

then $\{x \in E \mid f(x) = \infty\}$ has measure 0.

Proof. Let $F = \{x \in E \mid f(x) = \infty\}$, and $F_n = \{x \in E \mid f(x) > n\}$. Then $\forall n \in \mathbb{Z}^+$,

$$n\chi_F \leq f\chi_F \Longrightarrow n \cdot m(F) \leq \int_E f\chi_F \leq \int_E f < \infty.$$

Then $\forall n \in \mathbb{Z}^+$,

$$m(F) \leq \frac{1}{n} \int_E f \to 0.$$

Thus, $m(F) = 0$. □

12 Lecture 12. Lebesgue Integrals/Integrable Functions, Dominated Convergence Theorem

We've thus far defined the Lebesgue integral for nonnegative measurable functions. Now we will define the Lebesgue integral for a more general class of functions.

12.1 Lebesgue Integrable Functions

Definition 12.1. Let $E \subset \mathbb{R}$ be measurable. A function $f : E \to \mathbb{R}$ is Lebesgue integrable over E if

$$\int_E |f| < \infty.$$

Recall that we can write $f = f^+ - f^-$, so that $|f| = f^+ + f^-$. Therefore,

$$\int_E |f| = \int_E f^+ + \int_E f^-.$$

Thus, f is (Lebesgue) integrable iff f^+, f^- are integrable.

Definition 12.2. If $f : E \to \mathbb{R}$ is integrable, then the Lebesgue integral of f is

$$\int_E f := \int_E f^+ - \int_E f^-.$$

Theorem 12.3. *If $f, g : E \to \mathbb{R}$ are integrable,*

1. *$\forall c \in \mathbb{R}$, cf is integrable, and*
$$\int_E cf = c \int_E f,$$
2. *$f + g$ is integrable, and*
$$\int_E (f + g) = \int_E f + \int_E g.$$
3. *If A, B are disjoint measurable sets, then*
$$\int_{A \cup B} f = \int_A f + \int_B f.$$

Proof. We say property 1 is clear. For 2, note that $|f + g| \leq |f| + |g|$. Therefore,

$$\int_E |f + g| \leq \int_E |f| + |g| = \int_E |f| + \int_E |g| < \infty.$$

Thus, $f + g$ is integrable. Also,

$$f + g = f^+ + g^+ - (f^- + g^-) \implies (f + g)^+ + (f^- + g^-) = (f^+ + g^+) + (f + g)^-,$$

hence

$$\int_E (f + g)^+ + \int_E (f^- + g^-) = \int_E (f^+ + g^+) + \int_E (f + g)^-.$$

Hence

$$\int_E (f+g)^+ - \int_E (f+g)^- = \int_E (f^+ + g^+) - \int_E (f^- + g^-)$$
$$\implies \int_E f+g = \int_E f^+ + \int_E f^+ - \int_E f^- - \int_E g^-$$
$$= \int_E f + \int_E g.$$

Property 3 follows from 2 in the fact that $f\chi_{A\cup B} = f\chi_A + f\chi_B$, when $A \cap B = \emptyset$. ☐

Theorem 12.4. *Suppose $f, g : E \to \mathbb{R}$ are measurable.*

1. *If f is integrable, then $|\int_E f| \leq \int_E |f|$.*
2. *If g is integrable, and $f = g$ almost everywhere, then f is integrable, and $\int_E f = \int_E g$.*
3. *If f, g are integrable, and $f \leq g$ almost everywhere, then $\int_E f \leq \int_E g$.*

Proof. For 1,

$$\left|\int_E f\right| = \left|\int_E f^+ - \int_E f^-\right| \leq \int_E f^+ + \int_E f^- = \int_E f^+ + f^- = \int_E |f|.$$

For 2, we have $|f| = |g|$ almost everywhere. Thus

$$\int_E |f| = \int_E |g| < \infty.$$

Thus f is integrable. Moreover, $|f - g| = 0$ almost everywhere, thus

$$\left|\int_E f - \int_E g\right| = \left|\int_E (f-g)\right| \leq \int_E |f-g| = 0.$$

Thus, $\int_E f = \int_E g$. For 3, define a function

$$h(x) = \begin{cases} g(x) - f(x), & g(x) \geq f(x) \\ 0, & \text{otherwise.} \end{cases}$$

Then $h \in L^+(E)$, and $h = g - f$ almost everywhere, and $\int_E h < \infty$. Therefore,

$$0 \leq \int_E h^+ = \int_E h = \int_E g - f = \int_E g - \int_E f.$$

☐

12.2 Dominated Convergence Theorem

This is probably the most powerful theorem of the Lebesgue theory. Before we state this theorem, what are some functions that are Lebesgue integrable? What sets have finite measure? We know the measure of intervals are the lengths of those intervals, so any measurable sets that are contained in large intervals have finite measure. Compact sets, which we know are measureble because they are Borel sets. If we have simple functions which are nonzero only on sets that have finite measure, then we know they will be integrable. What about continuous functions on closed and bounded intervals $[a,b]$? If a function f is continuous on $[a,b]$, then its absolute value on $[a,b]$ is also continuous, hence its absolute value must be bounded by some constant. Therefore by the monotonicity of Lebesgue integrals, the integral of $\int_{[a,b]}|f| \leq \int_{[a,b]} C = Cm([a,b]) < \infty$. This shows that continuous functions on closed and bounded intervals are Lebesgue integrable. Let's prove something much stronger.

Theorem 12.5 (Lebesgue's dominated convergence theorem)**.** *Let $g : E \to [0,\infty]$ be a Lebesgue integrable function, and $\{f_n\}$ be a sequence of real-valued measurable functions such that*

1. $\forall n \in \mathbb{Z}^+$, $|f_n| \leq g$ *almost everywhere,*
2. $\exists f : E \to \mathbb{R}$ *such that $f_n \to f$ pointwise almost everywhere.*

Then

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

Proof. Since $\forall n, |f_n| \leq g$ a.e., $\forall n, f_n$ is integrable. Moreover, $f_n \to f$ a.e. implies that f is measurable and $|f| \leq g$ a.e., which implies that f is integrable. Since changing f_n for each n. Since changing f and f_n, $\forall n \in \mathbb{Z}^+$ on a set of measure zero does not affect the integrals, we can assume that the two given assumptions hold everywhere.

Now, note that $\forall n \in \mathbb{Z}^+$,

$$\left|\int_E f_n\right| \leq \int_E |f_n| \leq \int_E g.$$

Hence, $\{\int_E f_n\}$ is a bounded sequence in $\mathbb{R}$. Since $g \pm f_n \geq 0$, by Fatou's lemma,

$$\int_E g - f = \int_E \liminf_{n\to\infty}(g - f_n) \leq \liminf_{n\to\infty} \int_E g - f_n = \int_E g - \limsup_{n\to\infty} \int_E f_n.$$

(Recall that $\liminf_{n\to\infty}(-a_n) = -\limsup_{n\to\infty} a_n$.) Similarly,

$$\int_E g + f \leq \int_E g + \liminf_{n\to\infty} \int_E f_n.$$

Rearranging terms,

$$\begin{aligned}
\limsup_{n\to\infty} \int_E f_n &\leq \int_E g - \int_E g - f \\
&= \int_E f \\
&= \int_E g + f - \int_E g \\
&\leq \liminf_{n\to\infty} \int_E f_n.
\end{aligned}$$

Conclude that

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

□

I bet you thought in analysis that lim infs and lim sups would never be useful.

Prof. Casey Rodriguez

Theorem 12.6. *Suppose $a < b$, and $f \in C[a,b]$. Then*

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

That is, the Lebesgue integral of f is the same as its Riemann integral.

Proof. The same argument we used in the beginning of this chapter holds: $f \in C[a,b]$ implies that $|f| \in C[a,b]$ which implies that $\exists B \geq 0, |f| \leq B$ on $[a,b]$. Then

$$\int_{[a,b]} |f| \leq \int_{[a,b]} B = Bm(a,b) = B(b-a) < \infty.$$

Thus f is Lebesgue integrable. Consider that we can write

$$f^+ = \frac{f+|f|}{2}, \quad f^- = \frac{|f|-f}{2}.$$

By considering these separately, and showing that their Lebesgue and Riemann integrals agree, and using linearity, we may assume that $f \geq 0$. Let $x^n = \{x_0^n, x_1^n, \ldots, x_m^n\}$ be a sequence of partitions, where $x_0^n = a$, $x_m^n = b$, of $[a,b]$, such that $\|x^n\| := \max_{1\leq j \leq m_n} |x_j - x_{j-1}| \to 0$. Let $\xi_j^n \in [x_{j-1}^n, x_j^n]$, s.t.

$$\inf_{x \in [x_{j-1}^n, x_j^n]} f(x) = f(\xi_j^n).$$

Then by the theory of Riemann integration,

$$\lim_{n\to\infty} \sum_{j=1}^{m_n} f(\xi_j^n)(x_j^n - x_{j-1}^n) = \int_a^b f(x)\,dx.$$

Let $N = \bigcup_{n=1}^\infty x^n$. Then N is a countable union of finite sets, so N is countable, and $m(N) = 0$. Let

$$f_n = \sum_{j=1}^{m_n} f(\xi_j^n)\chi_{[x_{j-1}^n, x_j^n)} + 0\chi_{\{x_j^n\}}$$

be a nonnegative simple function. Note for all $n \in \mathbb{Z}^+$,

$$\begin{aligned}\int_{[a,b]} f_n &= \sum_{j=1}^{m_n} f(\xi_j^n) m([x_{j-1}^n, x_j^n]) \\ &= \sum_{j=1}^{m_n} f(\xi_j^n)(x_j^n - x_{j-1}^n).\end{aligned}$$

Then $\forall x \in [a,b] \setminus N$,

$$0 \leq f_n(x) \leq f(x).$$

Now we claim if $x \in [a,b] \setminus N$,

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

almost everywhere. Let $x \in [a,b] \setminus N$. Let $\varepsilon > 0$. Since f is continuous at x, $\exists \delta > 0$ such that if $|x-y| < \delta \implies |f(x) - f(y)| < \varepsilon$. Since $\|x^n\| = \max_{1\le j\le n}(x_j^n - x_{j-1}^n) \to 0$, $\exists M \in \mathbb{Z}^+$, such that $\forall n \geq M$,

$$\max_{1\le j\le n}(x_j^n - x_{j-1}^n) < \delta.$$

Let $n \geq M$. Then

$$\begin{aligned} f_n(x) &= \sum_{j=1}^{m_n} f(\xi_j^n)\chi_{[x_{j-1}^n, x_j^n]}(x) \\ &= f(\xi_k^n), \quad \text{for a unique } k \text{ s.t. } x \in (x_{k-1}^n, x_k). \end{aligned}$$

Then since $\xi_k^n \in [x_{k-1}^n, x_k^n]$, and $x_k^n - x_{k-1}^n < \delta$, we have that $|x - \xi_k^n| < \delta$, thus

$$|f(x) - f_n(x)| = |f(x) - f(\xi_k^n)| < \varepsilon.$$

Therefore, we've shown that $f_n \to f$ pointwise for $x \in [a,b] \setminus N$. Since $m(N) = 0$, we conclude that $f_n \to f$ pointwise a.e. and that $0 \leq f_n(x) \leq f(x)$ a.e., so by the DCT,

$$\int_{[a,b]} f = \lim_{n\to\infty}\int_{[a,b]} f_n = \lim_{n\to\infty}\sum_{j=1}^{m_n} f(\xi_j^n)(x_j^n - x_{j-1}^n) = \int_a^b f(x)\,dx.$$

□

The previous theorems we've proven easily imply corresponding statements for complex-valued integrable functions.

Definition 12.7. We say $f : E \to \mathbb{C}$ is Lebesgue integrable if $\int_E |f| < \infty$, with

$$\int_E f := \int_E \operatorname{Re} f + i \int_E \operatorname{Im} f.$$

Linearity of the Lebesgue integral, DCT, can be generalized relatively pain free to complex valued integrable functions. For example.

Theorem 12.8. *If $f : E \to \mathbb{C}$ is integrable, then*

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

Proof. This is clear if $\int_E f = 0$. Suppose $\int_E f \neq 0$. Let

$$\alpha = \frac{\overline{\int_E f}}{|\int_E f|}.$$

Then $|\alpha| = 1$, and

$$\left|\int_E f\right| = \alpha\int_E f = \int_E \alpha f = \operatorname{Re}\int_E \alpha f = \int_E \operatorname{Re}(\alpha f) \leq \int_E |\operatorname{Re}(\alpha f)| \leq \int_E |\alpha f| = \int_E |f|.$$

□

13 Lecture 13. Lp Space Theory

The whole point of our measure theoretic endeavour was to find a complete space of integrable function, whatever integrable meant, containing the space of continuous functions, with norm given by the integral of the pth power of the continuous function. We have achieved enough to continue our discussion of what is considered functional analysis. But there is more to be proven w.r.t. the Lebesgue theory. One can show that using the DCT that every Riemann integrable function on a closed and bounded interval is Lebesgue integrable, and that the Riemann integral is the Lebesgue integral. An equally important result is that a measurable function is Riemann integrable iff it is continuous almost everywhere. (Material to be covered in a measure theory course.)

13.1 Lp Spaces

Definition 13.1. If $F : E \to \mathbb{C}$ is measurable, and $1 \leq p < \infty$, we define

$$\|f\|_{L^p(E)} = \left(\int_E |f|^p \right)^{\frac{1}{p}}.$$

We define

$$\|f\|_{L^\infty(E)} = \inf\{M > 0 \mid m(\{x \in E \mid |f(x)| > M = 0\})\} =: \text{ess sup}_{x\in E}|f(x)|.$$

Theorem 13.2. *If $f : E \to \mathbb{C}$ is measurable, then*

$$|f(x)| \leq \|f\|_{L^\infty(E)}, \text{ a.e. on } E.$$

If $E = [a, b]$, and $f \in C[a, b]$, then

$$\|f\|_{L^\infty([a,b])} = \|f\|_\infty = \sup_{x\in[a,b]} |f(x)|.$$

Theorem 13.3 (Hölder's inequality)**.** *If $1 \leq p \leq \infty$, and $\frac{1}{p} + \frac{1}{q} = 1$, and $f, g : E \to \mathbb{C}$ are measurable, then*

$$\int_E |fg| \leq \|f\|_{L^p(E)} \|g\|_{L^q(E)}.$$

Theorem 13.4 (Minkowski's inequality)**.** *If $1 \leq p \leq \infty$, and $f, g : E \to \mathbb{C}$ are measurable, then*

$$\|f + g\|_{L^p(E)} \leq \|f\|_{L^p(E)} + \|g\|_{L^p(E)}.$$

In the future, we will denote $\|\cdot\|_{L^p(E)} = \|\cdot\|_p$. The traditional definition of L^p spaces is as follows, and it involves some traditional abuse of notation. But abuse of notation is tradition in math.

Definition 13.5. For $1 \leq p \leq \infty$, we define

$$L^p(E) = \{f : E \to \mathbb{C} \mid f \text{ is measurable, and } \|f\|_p < \infty\}.$$

This space is treated as a space of functions, and we talk about elements in $L^p(E)$ "functions." But this space is actually a space of equivalence classes. We consider two elements $f, g \in L^p(E)$ to be the same element, if $f = g$ almost everywhere. This allows us to consider the L^p norm as an actual norm. Otherwise, the L^p norm is only a seminorm. It is customary to never refer to elements of $L^p(E)$ as equivalence classes, instead calling them functions.

Theorem 13.6. *The set $L^p(E)$ with the natural scalar multiplication and addition operations is a vector space. Moreover, $\|\cdot\|_p$ is a norm on $L^p(E)$.*

Proof. When verifying $L^p(E)$ is a vector space, take care to remember that elements of $L^p(E)$ are equivalence classes. We'll skip the proof that $L^p(E)$ is a vector space, and verify $\|\cdot\|_p$ is a norm. The norm $\|\cdot\|_p$ is well defined on $L^p(E)$. If $f = g$ a.e., then

$$\int_E |f|^p = \int_E |g|^p.$$

Thus, if $[f], [g] \in L^p(E)$, $[f] = [g]$, then $\|[f]\|_p = \|[g]\|_p$. Now,

$$\int_E |f|^p = 0 \iff |f^p| = 0 \text{ a.e.} \iff f = 0 \text{ a.e.} \iff [f] = [0].$$

Homogeneity and the triangle inequality for $\|\cdot\|$ follow from its definition, and Minkowski's inequality. □

Let's talk about L^p spaces! How do we find functions in $L^p(E)$? Are these spaces even big enough to study?

Theorem 13.7. *Let $E \subset \mathbb{R}$ be measurable. Then $f \in L^p(E)$ iff*

$$\lim_{n\to\infty} \int_{[-n,n]\cap E} |f|^p < \infty.$$

Proof. Note that $\int_{[-\infty,\infty]\cap E} |f|^p = \int_E \chi_{[-\infty,\infty]} |f|^p$. Since

$$\chi_{[-1,1]}|f|^p \le \chi_{[-2,2]}|f|^p \le \dots,$$

on E, and $\forall x \in E, \lim_{n\to\infty} \chi_{[-n,n]}(x)|f(x)|^p = |f(x)|^p$, we have by the monotone convergence theorem,

$$\int_E |f|^p = \lim_{n\to\infty} \int_{[-n,n]\cap E} |f|^p.$$

□

Using this theorem, we can prove the following.

Example 13.8. If $f : \mathbb{R} \to \mathbb{C}$ is measurable, and $\exists C \ge 0$ and $q > 1$ such that for a.e. $x \in \mathbb{R}$,

$$|f(x)| \le C(1 + |x|)^{-q},$$

then $f \in L^p(\mathbb{R}), \forall p \ge 1$. Consider

$$\int_{[-n,n]} |f|^p \le \int_{[-n,n]} C(1 + |x|)^{-pq} = \int_{-n}^{n} C(1 + |x|)^{-pq}\, dx,$$

where the second equality holds as $C(1 + |x|)^{-pq}$ is continuous on a compact set. We are left to verify that as long as $q > 1$, the integral above is bounded by $C^p B(p)$, for some B depending on p.

I note that a lot of the theorems being stated are proven as homeworks. These are pretty important theorems, but unfortunately, I'm only taking notes on functional analysis to supplement my other courses, so I don't have time to prove these things independently.

Theorem 13.9. *Let $a < b$, $1 \le p < \infty$, $f \in L^p([a,b])$, and $\varepsilon > 0$. Then $\exists g \in C[a,b]$, such that $g(a) = g(b) = 0$, and*

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

That is, $C[a,b]$ is dense in $L^p([a,b])$.

Theorem 13.10 (Riesz-Fischer)**.** *For E measurable, and $1 \le p \le \infty$, $L^p(E)$ is Banach.*

Proof. We'll prove the $1 \le p < \infty$ case. The $p = \infty$ case is an assignment problem. A long time ago, we proved that a space is Banach iff all absolutely summable sequences in the space are summable. Suppose $\{f_k\}$ is a sequence in $L^p(E)$ such that

$$\sum_{k=1}^{\infty} \|f_k\|_p = M < \infty.$$

We want to show $\exists f \in L^p(E)$ such that

$$\sum_{k=1}^{n} f_k \to f, \text{ or } \lim_{n\to\infty} \left\| \sum_{k=1}^{n} f_k - f \right\|_p = 0.$$

Define $g_n : E \to [0, \infty)$, via

$$g_n(x) = \sum_{k=1}^{n} |f_k(x)|.$$

Then

$$\begin{aligned} \|g_n\|_p &= \left\| \sum_{k=1}^{n} |f_k| \right\|_p \\ &\le \sum_{k=1}^{n} \|f_k\|_p \le M < \infty. \end{aligned}$$

By Fatou's lemma,

$$\begin{aligned} \int_E \left(\sum_{k=1}^{\infty} |f_k| \right)^p &= \int_E \liminf_{n\to\infty} |g_n|^p \\ &\le \liminf_{n\to\infty} \int_E |g_n|^p \\ &\le M^p. \end{aligned}$$

Since it's integral is finite, $\sum_{k=1}^{\infty} |f_k(x)| < \infty$ for a.e. $x \in E$. Define

$$f(x) = \begin{cases} \sum_{k=1}^{\infty} f_k(x), & \text{if } \sum_{k=1}^{\infty} |f_k(x)| < \infty \\ 0 & \text{otherwise,} \end{cases}$$

and

$$g(x) = \begin{cases} \sum_{k=1}^{\infty} |f_k(x)|, & \sum_{k=1}^{\infty} |f_k(x)| < \infty \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\lim_{n\to\infty} \left| \sum_{k=1}^{n} f_k(x) - f(x) \right|^p = 0, \text{ a.e. on } E.$$

Also,

$$\left| \sum_{k=1}^{n} f_k(x) - f(x) \right|^p \leq |g(x)|^p, \text{ a.e. on } E.$$

Since $\|\sum_{k=1}^{n} |f_k|\|_p \leq M$, $\|g\|_p \leq M$. Thus,

$$\int_E |g|^p < \infty.$$

Moreover, $\|f\|_p \leq \|g\|_p \leq M$, i.e., $f \in L^p(E)$. By the DCT,

$$\lim_{n\to\infty} \int_E \left| \sum_{k=1}^{n} f_k - f \right|^p = 0.$$

That is,

$$\lim_{n\to\infty} \left\| \sum_{k=1}^{n} f_k - f \right\|_p^p = 0.$$

□

We use a lot of tools we've developed throughout this course. This is a pretty big result. In fact, the two previous theorems we've discussed show that the completion of $C[a, b]$ is $L^p([a, b])$.

13.2 Pre-Hilbert Spaces

And now, we are done with the theory of measure and integration, and we will now go back to our discussion of general functional analysis. A lot of functional analysis results have analogs in classical linear algebra. There are certain results which certainly don't have analogs. Alot of exact analogies will occur for Hilbert spaces, and certain operators on Hilbert spaces will be analogous to self-adjoint operators. And of course, in an applied sense, quantum mechanics take place in Hilbert spaces. The additional structure of a norm coming from an inner product gives us a lot more to say about Hilbert spaces.

Definition 13.11 (Pre-Hilbert spaces)**.** A pre-Hilbert space H is a vector space over $\mathbb{C}$ with a Hermitian inner product. That is, a map $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{C}$ s.t.

1. $\forall \lambda_1, \lambda_2 \in \mathbb{C}, v_1, v_2, w \in H$, $\langle \lambda_1 v_1 + \lambda_2 v_2, w \rangle = \lambda_1 \langle v_1, w \rangle + \lambda_2 \langle v_2, w \rangle$,
2. $\forall v, w \in H$, $\langle v, w \rangle = \overline{\langle w, v \rangle}$,
3. $\forall v \in H, \langle v, v \rangle \geq 0, \langle v, v \rangle = 0 \iff v = 0$.

We get that if $v \in H$ and $\langle v, w \rangle = 0, \forall w \in H$, then $v = 0$. Also, $\forall v, w \in H, \lambda \in \mathbb{C}$,

$$\langle v, \lambda w \rangle = \overline{\langle \lambda w, v \rangle} = \overline{\lambda \langle w, v \rangle} = \overline{\lambda} \langle v, w \rangle.$$

Definition 13.12. If H is a pre-Hilbert space, then we define $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$.

Theorem 13.13 (Cauchy-Bunyakovski-Schwarz)**.** *For all* $u, v \in H$*,* $|\langle u, v\rangle| \leq \|u\|\|v\|$*.*

Proof. Let $f(t) = \|u + tv\|^2 \geq 0$. Then

$$\begin{aligned} f(t) &= \langle u, u\rangle + t^2\langle v, v\rangle + t\langle u, v\rangle + t\langle v, u\rangle \\ &= \|u\|^2 + t^2\|v\|^2 + 2t \operatorname{Re}\langle u, v\rangle. \end{aligned}$$

The minimum of f is greater or equal to 0, and

$$f'(t_{\min}) = 0 \iff t_{\min} = \frac{-\operatorname{Re}\langle u, v\rangle}{\|v\|^2},$$

thus

$$0 \leq f(t_{\min}) = \|u\|^2 - \frac{(\operatorname{Re}\langle u, v\rangle)^2}{\|v\|^2},$$

therefore,

$$|\operatorname{Re}\langle u, v\rangle| \leq \|u\|\|v\|.$$

If $\langle u, v\rangle = 0$, we are done, so suppose $\langle u, v\rangle \neq 0$. Let

$$\lambda = \frac{\overline{\langle u, v\rangle}}{|\langle u, v\rangle|}.$$

Then $|\lambda| = 1$, and

$$\begin{aligned} |\langle u, v\rangle| &= \lambda\langle u, v\rangle \\ &= \langle \lambda u, v\rangle \\ &= \operatorname{Re}\langle \lambda u, v\rangle \\ &\leq \|\lambda u\|\|v\| \\ &= \|u\|\|v\|. \end{aligned}$$

□

We're going to use the Cauchy-Schwarz inequality to prove that this norm we've been talking about actually is a norm. Next time, we'll prove that there are really only two types of reasonable or interesting Hilbert space. The first type are finite-dimensional, think $\mathbb{R}^n$, $\mathbb{C}^n$. The second type is $\ell^2(E)$. We'll show that every separable Hilbert space has a countable orthonormal basis.

14 Lecture 14. Basic Hilbert Space Theory

Let's prove the norm we've defined on pre-Hilbert spaces is a norm.

Theorem 14.1. *If* H *is a pre-Hilbert space, then* $\|\cdot\|$ *is a norm on* H*.*

Proof. Note $\|v\| = - \iff \langle v, v\rangle = 0 \iff v = 0$. If $\lambda \in \mathbb{C}$, and $v \in H$,

$$\langle \lambda v, \lambda v\rangle = \lambda\overline{\lambda}\langle v, v\rangle = |\lambda|^2\langle v, v\rangle \implies \|\lambda v\| = |\lambda|\|v\|.$$

Let $u, v \in H$, then

$$\begin{aligned}
\|u+v\|^2 &= \langle u+v, u+v\rangle \\
&= \|u\|^2 + \|v\|^2 + 2\operatorname{Re}\langle u, v\rangle \\
&\le \|u\|^2 + \|v\|^2 + 2|\operatorname{Re}\langle u, v\rangle| \\
&\le \|u\|^2 + \|v\|^2 + 2\|u\|\|v\| \\
&= (\|u\| + \|v\|)^2 \\
\implies \|u+v\| &\le \|u\| + \|v\|.
\end{aligned}$$

□

Theorem 14.2 (Continuity of $\langle\cdot,\cdot\rangle$)**.** *If $u_n \to u$ and $v_n \to v$ in a pre-Hilbert space with the norm $\|\cdot\| = \sqrt{\|\cdot,\cdot\|}$, then*

$$\langle u_n, v_n\rangle \to \langle u, v\rangle.$$

Proof. If $u_n \to u$ and $v_n \to v$, i.e. $\|u_n - u\| \to 0$ and $\|v_n - v\| \to 0$ as $n \to \infty$, then

$$\begin{aligned}
|\langle u_n, v_n\rangle - \langle u, v\rangle| &= |\langle u_n - u, v_n\rangle + \langle u, v_n - v\rangle| \\
&\le |\langle u_n - u, v_n\rangle| + |\langle u, v_n - v\rangle| \\
&\le \|u_n - u\|\|v_n\| + \|u\|\|v_n - v\| \\
&\le \|u_n - u\| \sup_n \|v_n\| + \|v_n - v\|\|u\| \\
\implies |\langle u_n, v_n\rangle - \langle u, v\rangle| &\to 0,
\end{aligned}$$

by the squeeze theorem. □

14.1 Hilbert Spaces

Definition 14.3. A Hilbert space H is a pre-Hilbert space which is complete w.r.t. its norm $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$.

Example 14.4. We mentioned earlier that there are really only two reasonable forms of Hilbert spaces. One classically studied Hilbert space is $\mathbb{C}^n$ with the conjugate dot product. The other example we have is $\ell^2 = \{a = \{a_k\} \mid a_k \in \mathbb{C}, \sum_{k=1}^{\infty}|a_k|^2 < \infty\}$, with the inner product

$$\langle a, b\rangle = \sum_{k=1}^{\infty} a_k \overline{b_k}.$$

Then the norm we get from this inner product is

$$\left(\sum_{k=1}^{\infty}|a_k|^2\right)^{1/2} =: \|a\|_{\ell^2}.$$

We will soon show that every separable Hilbert space can be mapped isometrically to either $\mathbb{C}^n$ or ℓ^2.

Example 14.5. If $E \subset \mathbb{R}$ is measurable, then $L^2(E)$ is a Hilbert space, with inner product

$$\langle f, g\rangle = \int_E f\overline{g}.$$

Is there any inner product that exists for other general ℓ^p, L^p space, inducing the ℓ^p, L^p norms? In other words, are any other ℓ^p, L^p spaces Hilbert? The answer is no. The easiest way to determine a Hilbert space is to first define an inner product, then define the norm following the inner product. But if we are given just a norm on a space, is there a way to find the corresponding inner product?

Theorem 14.6 (Parallelogram law)**.** *If H is a pre-Hilbert space, then $\forall u, v \in H$,*

$$\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2).$$

We call this equality the parallelogram law, and moreover, if H is a normed space satisfying the parallelogram law, then H is a pre-Hilbert space.

Using this theorem we can verify that the only ℓ^p, L^p spaces which are Hilbert spaces are $p = 2$. The theorem is not terrible to prove, typically presented in a linear algebra course, and requiring basic analysis. Funnily enough, each linear algebra instructor I've had has either assigned this as a homework problem, or told us to consult Stack Exchange for a proof (even citing the exact post id: 21792.)

14.2 Orthonormal Sets

Definition 14.7. If H is a pre-Hilbert space, $u, v \in H$ are orthogonal if $\langle u, v\rangle = 0$ written sometimes $(u \perp v)$.

Definition 14.8. If H is a pre-Hilbert space, a subset $\{e_\lambda\}_{\lambda\in\Lambda} \subset H$ is orthonormal if $\forall \lambda \in \Lambda$, $\|e_\lambda\| = 1$ and $\lambda_1, \lambda_2 \in \Lambda, \lambda_1 \neq \lambda_2 \implies e_{\lambda_1} \perp e_{\lambda_2}$.

Note that although we are mainly concerned with countable or finite sets in this course, when we write notation such as $\lambda \in \Lambda$ we mean that these notions can be extended to uncountable sets.

Example 14.9. The set

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} \subset \mathbb{C}^3$$

is an orthonormal set. The sequence $\{e_n\}$ defined by

$$e_n = (0, 0, \ldots, 0, \overbrace{1}^{\text{n}th \text{ entry}}, 0, \ldots)$$

is an orthonormal set of ℓ^2.

Example 14.10. Consider $\frac{1}{\sqrt{2\pi}} e^{inx} \in L^2([-\pi, \pi])$. Then

$$\left\{ \frac{1}{\sqrt{2\pi}} e^{inx} \right\}_{n\in\mathbb{Z}^+} \subset L^2([-\pi,\pi])$$

is an orthonormal set. Consider for $m \neq n$,

$$\begin{aligned} \int_{-\pi}^{\pi} \frac{1}{\sqrt{2\pi}} e^{-inx} \overline{\frac{1}{\sqrt{2\pi}} e^{imx}}\, dx &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(m-n)x}\, dx \\ &= \frac{1}{2\pi} \frac{e^{i(m-n)x}}{i(m-n)} \bigg|_{-\pi}^{\pi} = 0. \end{aligned}$$

When $m = n$, we can also see this quantity equals 1.

As an aside, so long as we aren't working in L^∞, we can easily prove that there is a countably dense subset of L^p, meaning that L^p is a separable space. There was an assignment of this course where it was proven that continuous functions are dense in L^p. Then by the Weierstrauss approximation theorem, the polynomials are dense in the continuous functions. We also know that polynomials of rational coefficients are dense in the set of polynomials. And we know that this set is countable. The $\ell^p, p < \infty$ spaces are also separable. We can prove that the set of sequences which terminate after some finite index is dense in ℓ^p, in other words, sequences which are 0 identically after some index n. We can approximate those sequences by sequences of rational numbers which terminate after some finite index. This set of sequences is countable.

Theorem 14.11 (Bessel)**.** *If $\{e_n\}$ is a countable (finite or infinite) orthonormal subset of a pre-Hilbert space H, then $\forall u \in H$,*

$$\sum_{n=1}^{\infty} |\langle u, e_n\rangle|^2 \leq \|u\|^2.$$

Proof. In the finite case, suppose $\{e_1, e_2, \dots, e_n\}$ is an orthonormal subset of H. Then

$$\begin{aligned}
\sum_{n=1}^{N} \|\langle u, e_n\rangle e_n\|^2 &= \left\langle \sum_{n=1}^{N} \langle u, e_n\rangle e_n, \sum_{m=1}^{N} \langle u, e_m\rangle e_m \right\rangle \\
&= \sum_{n,m=1}^{N} \langle u, e_n\rangle \overline{\langle u, e_m\rangle} \langle e_n, e_m\rangle \\
&= \sum_{n=1}^{N} \langle u, e_n\rangle \overline{\langle u, e_n\rangle} \\
&= \sum_{n=1}^{N} |\langle u, e_n\rangle|^2.
\end{aligned}$$

Also,

$$\begin{aligned}
\left\langle u, \sum_{n=1}^{N} \langle u, e_n\rangle e_n \right\rangle &= \sum_{n=1}^{N} \overline{\langle u, e_n\rangle} \langle u_n, e_n\rangle \\
&= \sum_{n=1}^{N} |\langle u, e_n\rangle|^2.
\end{aligned}$$

Thus,

$$\begin{aligned}
0 &\leq \left\| u - \sum_{n=1}^{N} \langle u, e_n\rangle e_n \right\|^2 \\
&\leq \|u\|^2 + \left\| \sum_{n=1}^{N} \langle u, e_n\rangle e_n \right\|^2 - 2\operatorname{Re} \left\langle u, \sum_{n=1}^{N} \langle u, e_n\rangle e_n \right\rangle \\
&= \|u\|^2 - \sum_{n=1}^{N} |\langle u, e_n\rangle|^2.
\end{aligned}$$

The infinite case follows from the finite case by letting $N \to \infty$. Suppose $\{e_n\}$ is a countably infinite orthonormal subset of H. Then $\forall n \in \mathbb{Z}^+$,

$$\sum_{n=1}^{N} |\langle u, e_n\rangle|^2 \leq \|u\|^2 \implies \sum_{n=1}^{\infty} |\langle u, e_n\rangle|^2 \leq \|u\|^2$$

□

14.3 Gram Schmidt and Maximal Orthonormal Sets

Definition 14.12. An orthonormal subset $\{e_\lambda\}_{\lambda\in\Lambda}$ of a pre-Hilbert space H is maximal if $u \in H$ and $\langle u, e_\lambda\rangle = 0, \forall \lambda \in \Lambda \implies u = 0$.

Example 14.13. The set

$$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

is maximal in $\mathbb{C}^2$, and the set

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is not maximal in $\mathbb{C}^3$ since $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ is orthogonal to this set. The sequence $\{e_n\}$ descibed in Example 14.9 is maximal in ℓ^2. We will see shortly that these maximal sets serve the same roles as orthonormal bases in linear algebra. We won't see them strictly as Hamel bases, but possibly infinite analogs. In most cases, this is just as good as a Hamel basis.

Theorem 14.14. *Every nontrivial pre-Hilbert space has a maximal orthonormal subset.*

The proof of this uses Zorn's lemma. We won't present this proof, but we will prove something just as useful. Zorn's lemma is equivalent to the axiom of choice, so maybe we want something we can make sense of without the AoC.

Theorem 14.15. *Every nontrivial separable pre-Hilbert space has a countable maximal orthonormal subset.*

Proof. Let $\{v_j\}$ be a countably dense subset of H such that WLOG $\|v_1\| \neq 0$. We claim that $\forall n \in \mathbb{Z}^+$, there exists $m(n) \leq n$ and an orthonormal subset $\{e_1, \ldots, e_{m(n)}\}$ such that

1. $\operatorname{span}\{e_1, \ldots, e_{m(n)}\} = \operatorname{span}\{v_1, \ldots, v_m\}$,
2. $\{e_1, \ldots, e_{m(n)}\} = \{e_1, \ldots, e_{m(n-1)}\} \cup \begin{cases} \emptyset, & \text{if } v_n \in \operatorname{span}\{v_1, \ldots, v_{n-1}\} \\ e_{m(n)}, & \text{otherwise.} \end{cases}$

By induction, in the case $n = 1$, $e_1 = v_1/\|v_1\|$. Suppose that we can find $\{e_1, \ldots, e_{m(n-1)}\}$ satisfying our claims 1, 2. If $v_n \in \operatorname{span}\{v_1, \ldots, v_{n-1}\}$, then $\operatorname{span}\{e_1, \ldots, e_{m(n)}\} = \operatorname{span}\{v_1, \ldots, v_{n-1}\} = \operatorname{span}\{v_1, \ldots, v_n\}$. Suppose that $v_n \notin \operatorname{span}\{v_1, \ldots, v_{n-1}\}$. Define

$$w_n = v_n - \sum_{j=1}^{m(n-1)} \langle v_n, e_j\rangle e_j.$$

Note that $w_n \neq 0$, otherwise $v_n \in \operatorname{span}\{v_1, \ldots, v_{n-1}\}$. Define $e_{m(n)} = \frac{w_n}{\|w_n\|}$. Then $\|e_{m(n)}\| = 1$, and for $1 \leq \ell \leq m(n-1)$

$$\begin{aligned}\langle e_{m(n)}, e_j \rangle &= \frac{1}{\|w_n\|} \left\langle v_n - \sum_{j=1}^{m(n-1)} \langle v_n, e_j \rangle e_j, e_\ell \right\rangle \\ &= \frac{1}{\|w_n\|} \left(\langle v_n, e_\ell \rangle - \langle v_n, e_\ell \rangle \right) = 0.\end{aligned}$$

Thus, $\{e_1, \ldots, e_{m(n)}\}$ completes our inductive step. Now, consider

$$S := \bigcup_{n=1}^{\infty} \{e_1, \ldots, e_{n(n)}\}.$$

(We will assume that S is countably infinite, but the argument is the same for when S is finite.) We see S is an orthonormal set of H. We will now show S is maximal. Suppose $u \in H$, and $\forall \ell \in \mathbb{Z}^+, \langle u, e_\ell \rangle = 0$. Since $\{v_j\}$ is a dense subset of H, there exists some $\{v_{j(k)}\}_{k=1}^{\infty}$ such that $v_{j(k)} \to u$ as $k \to \infty$. By property 1 of our e_j's, $v_{j(k)} \in \operatorname{span}\{e_1, \ldots, e_{m(j(k))}\}$, we get as a result of Bessel's inequality,

$$\begin{aligned}\|v_{j(k)}\|^2 &= \sum_{\ell=1}^{m(j(k))} |\langle v_{j(k)}, e_\ell \rangle|^2 \\ &= \sum_{\ell=1}^{m(j(k))} |\langle v_{j(k)} - u, e_\ell \rangle|^2 \\ &\leq \|v_{j(k)} - u\|^2 \to 0.\end{aligned}$$

Therefore, $\|v_{j(k)}\| \to 0$, so $u = 0$. □

15 Lecture 15. Orthonormal Bases and Fourier Series

Turns out, we have a special name for countable maximal orthonormal subsets.

Definition 15.1. Let H be a Hilbert space. An orthonormal basis of H is a countable maximal orthonormal subset $\{e_n\}$ of H.

We've dealt with Hamel bases: bases which you can write every element in a vector space as a finite linear combination of elements from the basis. In the setting of Hilbert spaces, in what sense are orthonormal bases playing the role of a basis?

Theorem 15.2. *If $\{e_n\}$ is an orthonormal basis in a Hilbert space H, then $\forall u \in H$,*

$$\sum_{n=1}^{m} \langle u, e_n \rangle e_n \to u,$$

as $m \to \infty$. That is,

$$u = \sum_{n=1}^{\infty} \langle u, e_n \rangle e_n.$$

We sometimes refer to this as a Fourier-Bessel series.

Proof. First, consider that $\{\sum_{n=1}^{m}\langle u, e_n\rangle e_n\}_{m=1}^{\infty}$ is Cauchy. Let $\varepsilon > 0$. By Bessel's inequality,

$$\sum_{n=1}^{\infty}|\langle u, e_n\rangle|^2 \leq \|u\|^2 < \infty.$$

Thus, $\exists M \in \mathbb{Z}^+, \forall N \geq M$,

$$\sum_{n=N+1}^{\infty} |\langle u, e_n\rangle|^2 < \varepsilon^2.$$

Thus, $\forall m > \ell \geq M$,

$$\left\|\sum_{n=1}^{m}\langle u, e_n\rangle e_n - \sum_{n=1}^{\ell}\langle u, e_n\rangle e_n\right\| = \sum_{n=l+1}^{m} |\langle u, e_n\rangle|^2 \leq \sum_{n=\ell+1}^{\infty} |\langle u, e_n\rangle|^2 < \varepsilon^2.$$

Since H is complete, there exists $u^* \in H$ such that $u^* = \lim_{m\to\infty}\sum_{n=1}^{m}\langle u, e_n\rangle e_n$. By continuity of $\langle\cdot,\cdot\rangle$, $\forall \ell \in \mathbb{Z}^+$,

$$\begin{aligned}
\langle u - u*, e_\ell\rangle &= \lim_{m\to\infty}\left\langle u - \sum_{n=1}^{m}\langle u, e_n\rangle e_n, e_\ell\right\rangle \\
&= \lim_{m\to\infty}\left(\langle u, e_\ell\rangle - \sum_{n=1}^{m}\langle u, e_n\rangle\langle e_n, e_\ell\rangle\right) \\
&= \lim_{m\to\infty}\left(\langle u, e_\ell\rangle - \langle u, e_\ell\rangle\right) \\
&= 0.
\end{aligned}$$

Since $\{e_n\}$ is a maximal countable orthonormal basis, $u - u^* = 0$, hence $u = u^*$. □

We've shown that if H is separable, then it has an orthonormal basis, and we also have the other direction holding as well.

Theorem 15.3. *If H has an orthonormal basis, then H is separable.*

Proof. Suppose $\{e_n\}$ is an orthonormal basis for H. Then

$$S = \bigcup_{m\in\mathbb{Z}^+}\left\{\sum_{n=1}^{m} q_n e_n \mid q_1, \ldots, q_m \in \mathbb{Q} + i\mathbb{Q}\right\}$$

is a countable set. (The rationals are countable, and cartesian products of countable sets are countable. Unions of countable sets are countable.) By the previous theorem, S is dense in H. (Take an element in $h \in H$. Consider its Fourier-Bessel series. Truncate the series, and its coefficients will be some numbers in $\mathbb{C}$. The rationals are dense in $\mathbb{R}$, so we can approximate the coefficients of the truncated series by rationals.) □

So if H is a Hilbert space, H is separable iff it has an orthonormal basis. With all of this new weaponry at hand, we can view Bessel's inequality as stating that the sum of the coefficients appearing in an element $u \in H$'s Fourier-Bessel series squared is always less than or equal to the norm of that element u squared. This result holds for any orthonormal subset, but in fact, we get equality when we use an orthonormal basis.

Theorem 15.4. *If H is a Hilbert space, and $\{e_n\}$ is a countable orthonormal basis (possibly finite), then $\forall u \in H$,*

$$\sum_n |\langle u, e_n \rangle|^2 = \|u\|^2.$$

Sometimes, this is referred to as Parseval's identity.

Proof. We have $u = \sum_n \langle u, e_n \rangle e_n$, thus

$$\begin{aligned}
\|u\|^2 &= \lim_{m\to\infty} \left\langle \sum_{n=1}^{m} \langle u, e_n \rangle e_n, \sum_{\ell=1}^{m} \langle u, e_\ell \rangle e_\ell \right\rangle \\
&= \lim_{m\to\infty} \sum_{n,\ell=1}^{m} \langle u, e_n \rangle \overline{\langle u, e_\ell \rangle} \langle e_n, e_\ell \rangle \\
&= \lim_{m\to\infty} \sum_{n=1}^{m} |\langle u, e_n \rangle|^2 = \sum_n |\langle u, e_n \rangle|^2.
\end{aligned}$$

□

These previous theorems give us a way to identify every separable Hilbert space. In fact, if H is a finite dimensional separable Hilbert space, it's quite easy to show that H is isomorphic, isometrically to $\mathbb{C}^n$. We will state the following about infinite dimensional separable Hilbert spaces.

Theorem 15.5. *If H is a infinite dimensional separable Hilbert space, then H is isometrically isomorphic to ℓ^2. Meaning there exists a bijective linear operator $T : H \to \ell^2$ such that $\forall u, v \in H$,*

$$\|Tu\|_{\ell^2} = \|u\|_H.$$

(Thus, $\langle Tv, Tv \rangle_{\ell^2} = \langle u, v \rangle_H$).

Proof. Since H is a separable Hilbert space, it has an orthonormal basis $\{e_n\}$, $\forall u \in H$,

$$u = \sum_{n=1}^{\infty} \langle u, e_n \rangle e_n,$$

with

$$\|u\| = \left(\sum_{n=1}^{\infty} |\langle u, e_n \rangle|^2 \right)^{1/2}.$$

Define

$$Tu = \{\langle u, e_n \rangle\}_{n=1}^{\infty} \in \ell^2.$$

Then T does the job, and the rest of the proof isn't too bad. □

15.1 Introductory Fourier Analysis

Let's take a pause from the general theory of functional analysis, and talk a little about Fourier series. The reason why alot of integration theory was developed was because of Fourier analysis.

Theorem 15.6. *The subset of functions*

$$\left\{ \frac{e^{inx}}{\sqrt{2\pi}} \right\}_{n\in\mathbb{Z}}$$

is an orthonormal subset of $L^2([-\pi, \pi])$.

Proof. Consider that

$$\begin{aligned}
\langle e^{inx}, e^{imx} \rangle &= \int_{-\pi}^{\pi} e^{inx}\overline{e^{imx}}\, dx \\
&= \int_{-\pi}^{\pi} e^{inx} e^{-imx}\, dx \\
&= \int_{-\pi}^{\pi} e^{i(n-m)x}\, dx \\
&= \begin{cases} 2\pi & n = m \\ \frac{e^{i(n-m)x}}{i(n-m)}\Big|_{-\pi}^{\pi} = 0 & n \neq m. \end{cases}
\end{aligned}$$

(Recall that $e^{ix} = \cos t + i \sin t$, which is 2π-periodic.) ☐

Definition 15.7. Let $f \in L^2([-\pi, \pi])$. The nth Fourier coefficient of f is

$$\hat{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int}\, dt.$$

The nth partial Fourier sum of f is

$$S_N f(x) = \sum_{|n| \leq N, n \in \mathbb{Z}} \hat{f}(n) e^{inx} = \sum_{|n| \leq N, n \in \mathbb{Z}} \left\langle f, \frac{e^{inx}}{\sqrt{2\pi}} \right\rangle \frac{e^{inx}}{\sqrt{2\pi}}.$$

Definition 15.8. The Fourier series of f is the formal series

$$\sum_{n \in \mathbb{Z}} \hat{f}(x) e^{inx}.$$

When Fourier was studying heat conduction, he claimed that every function can be expanded into a Fourier series, at the time, sums of sines and cosines. People back then were like, "Man, what are you talking about, not every function is periodic." But now we are looking at an orthonormal subset of L^2. Now we are tasked with answering, for all $f \in L^2([-\pi, \pi])$,

$$f(x) = \sum_n \hat{f}(x) e^{inx}?$$

That is, does

$$\|f - S_N f\|_2 = \left(\int_{-\pi}^{\pi} |f(x) - S_n f(x)|^2\, dx \right) \to 0, \text{ as } N \to \infty?$$

Also equivalently, is $\left\{ \frac{e^{inx}}{\sqrt{2\pi}} \right\}_{n \in \mathbb{Z}}$ maximal in $L^2([-\pi, \pi])$? Also equivalently, if

$$\hat{f}(n) = 0, \forall n \in \mathbb{Z} \implies f = 0 \tag{Fejer}$$

The answer is yes, but this is nontrivial. We will aim to show that Eq. (Fejer) holds. This result is proven by something usually referred to as Fejer's method, which we will present as follows.

Theorem 15.9. *For all $f \in L^2([-\pi,\pi])$, for all $n \in \mathbb{Z}_{\geq 0}$,*

$$S_N f(x) = \int_{-\pi}^{\pi} D_n(x-t) f(t)\, dt, \tag{*}$$

where

$$D_N(x) = \begin{cases} \dfrac{2N+1}{2\pi}, & x = 0 \\ \\ \dfrac{\sin(N+1/2)x}{2\pi \sin(x/2)}, & x \neq 0. \end{cases}$$

This function D_N is referred to as the Direchlet kernel, and the integration in $(*)$ is Lebesgue.

Proof. If $f \in L^2([-\pi,\pi])$,

$$\begin{aligned} S_N f(x) &= \sum_{|n| \leq N, n \in \mathbb{Z}} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int}\, dt \right) e^{inx} \\ &= \int_{-\pi}^{\pi} f(t) \underbrace{\left(\frac{1}{2\pi} \sum_{|n| \leq N, n \in \mathbb{Z}} e^{in(x-t)} \right)}_{D_N(x-t)} dt. \end{aligned}$$

Then

$$\begin{aligned} D_N(x) &= \frac{1}{2\pi} \sum_{n=-N}^{N} e^{inx} \\ &= \frac{1}{2\pi} e^{-iNx} \sum_{n=0}^{2N} e^{inx} \qquad (**) \\ &= \frac{1}{2\pi} e^{-iNx} \left(\frac{1 - e^{i(2N+1)x}}{1 - e^{ix}} \right) \\ &= \frac{1}{2\pi} \frac{e^{i(N+(1/2)x)} - e^{-i(N+(1/2))x}}{e^{i(x/2)} - e^{-i(x/2)}} \\ &= \frac{1}{2\pi} \frac{2i \sin(N+1/2)x}{2i \sin(x/2)} \\ &= \frac{1}{2\pi} \frac{\sin(N+1/2)x}{\sin(x/2)}, \end{aligned}$$

when $x \neq 0$. If $x = 0$, we get our desired result in $(**)$. □

Definition 15.10. If $f \in L^2([-\pi,\pi])$, we define the Nth Cesaro-Fourier mean of f to be

$$\sigma_N f(x) = \frac{1}{N+1} \sum_{k=0}^{N} S_k f(x).$$

Now instead of proving Eq. (Fejer), what if we were able to prove that $\sigma_N f(x) \to 0$? In analysis, if we have a sequence of numbers, we can define Cesaro means of that sequence, which behave nicer than the original sequence, but do not lose any information about that sequence. In fact, if a sequence of numbers converges, then its sequence of Cesaro means should converge. We get the added benefit of nicer convergence results. For example, in the real number case, consider the sequence $\{1, -1, 1, -1, \dots\}$, which does not converge, but its Cesaro means do converge: $\{1, 0, 1/3, 0, 1/5, \dots\}$, which converge to 0. We expect that $\sigma_N f(x)$ has better convergence properties than $S_N f(x)$. Mainly we hope that the Cesaro-Fourier means of f can show that the Fourier series of f converge to f in L^2. Our goal for next lecture is to show that

$$\|\sigma_N f - f\|_2 \xrightarrow{N\to\infty} 0.$$

If we are able to do this, then $0 \to f$, so f is 0, our goal in Eq. (Fejer).

16 Lecture 16. Fejer's Theorem and Convergence of Fourier Series

Recall that for $f \in L^2([-\pi, \pi])$, we define

$$\hat{f}(n) = \int_{-\pi}^{\pi} f(t) e^{-int}\, dt.$$

We also had that

$$S_N f(x) = \sum_{i=-N}^{N} \hat{f}(n) e^{inx}.$$

So our question is, $\forall f \in L^2([-\pi, \pi])$, does

$$\lim_{N\to\infty} \|f - S_N f\|_2 = 0?$$

Based on what we've learned about Hilbert spaces, this question is equivalent to asking, if

$$f \in L^2([-\pi, \pi]), \text{ and } \hat{f}(n) = 0, \forall n \in \mathbb{Z},$$

then does $f = 0$? We defined

$$\sigma_N f(x) = \frac{1}{N+1} \sum_{k=0}^{N} S_k f(x),$$

the Nth Cesaro-Fourier mean of f. What we're going to show is that if $f \in L^2([-\pi, \pi])$, then

$$\|\sigma_N f - f\|_2 \xrightarrow{N\to\infty} 0.$$

16.1 Fejer's Method

Theorem 16.1. *For all $f \in L^2([-\pi, \pi])$,*

$$\sigma_N f(x) = \int_{-\pi}^{\pi} K_N(x-t) f(t)\, dt,$$

where

$$K_N(x) = \begin{cases} \frac{N+1}{2\pi}, & x = 0 \\ \frac{1}{2\pi(N+1)} \left[\frac{\sin \frac{N+1}{2} x}{\sin \frac{x}{2}}\right]^2 & x \neq 0. \end{cases}$$

We call K_N Fejer's kernel. Moreover,

1. *$K_N(x) \geq 0$, $K_N(x) = K_N(-x)$, K_N is 2π-periodic.*
2. *The integral $\int_{-\pi}^{\pi} K_N(t)\,dt = 1$.*
3. *If $\delta \in (0, \pi]$, then $\forall x$, such that $\delta \leq |x| \leq \pi$, then*

$$|K_N(x)| \leq \frac{1}{2\pi(N+1)\sin^2 \frac{\delta}{2}}.$$

Proof. Last time we proved that

$$S_K f(x) = \int_{-\pi}^{\pi} D_K(x-t) f(t)\,dt,$$

where

$$D_K(t) = \begin{cases} \frac{2K+1}{2\pi}, & t = 0 \\ \frac{1}{2\pi}\frac{\sin(K+\frac{1}{2})t}{\sin\frac{t}{2}}, & t \neq 0. \end{cases}$$

Then

$$\begin{aligned} \sigma_N f(x) &= \frac{1}{N+1}\sum_{k=0}^{N} S_K f(x) \\ &= \int_{-\pi}^{\pi} \underbrace{\frac{1}{N+1}\sum_{k=0}^{N} D_k(x-t)}_{K_N(x-t)} f(t)\,dt. \end{aligned}$$

(In the case that $x = 0$, a simple calculation verifies the claim in the brackets.) Assume that $x \neq 0$. Then

$$\begin{aligned} K_N(x) &= \frac{1}{N+1}\sum_{k=0}^{N} D_k(x) \\ &= \frac{1}{2\pi(N+1)}\frac{1}{2(\sin\frac{x}{2})^2}\sum_{k=0}^{N} 2\sin\frac{x}{2}\sin\left(N+\frac{1}{2}\right)x \\ &= \frac{1}{2\pi(N+1)}\frac{1}{2\left(\sin\frac{x}{2}\right)^2}\sum_{k=0}^{n}(\cos(kx) - \cos((k+1)x)) \\ &= \frac{1}{2\pi(N+1)}\frac{1}{2\sin^2\frac{x}{2}}\left[(\cos 0x - \cos 1x) + (\cos 1x - \cos 2x) + \ldots + (\cos Nx - \cos{(N+1)x}\right] \\ &= \frac{1}{2\pi(N+1)}\frac{1}{2\sin^2\frac{x}{2}}\frac{1-\cos(N+1x)}{2} \\ &= \frac{1}{2\pi(N+1)}\frac{\sin^2\left(\frac{N+1}{2}x\right)}{\sin^2\frac{x}{2}}. \end{aligned}$$

The properties in 1 follow immediately. For 2, note

$$\int_{-\pi}^{\pi} D_k(t)\,dt = \int_{-\pi}^{\pi}\sum_{n=-k}^{k} e^{int}\,dt = 1.$$

Then

$$\int_{-\pi}^{\pi} K_N(t)\,dt = \frac{1}{N+1}\sum_{k=0}^{N}\int_{-\pi}^{\pi} D_k(t)\,dt = 1.$$

For 3, let $\delta \in (0,\pi]$. Then $\sin^2\frac{x}{2}$ is even and increasing on $[0,\pi]$. Thus, if $\delta \le |x| \le \pi$, then

$$\sin^2\frac{x}{2} \ge \sin^2\frac{\delta}{2}.$$

Then

$$K_N(x) \le \frac{1}{2\pi(N+1)}\frac{\sin^2\frac{N+1}{2}x}{\sin^2\frac{\delta}{2}} \le \frac{1}{2\pi(N+1)\sin^2\frac{\delta}{2}}.$$

□

Theorem 16.2 (Fejer)**.** *If $f \in C([-\pi,\pi])$ and 2π-periodic, that is $f(\pi) = f(-\pi)$, then*

$$\sigma_N f \to f \text{ uniformly on } [-\pi,\pi].$$

Proof. First, we extend f by periodicity to all of $\mathbb{R}$. Then $f \in C(\mathbb{R})$ is 2π-periodic, so f is uniformly continuous and bounded:

$$\|f\|_\infty = \sup_{x\in\mathbb{R}}|f(x)| = \sup_{x\in[-\pi,\pi]}|f(x)| < \infty.$$

Let $\varepsilon > 0$, then $\exists\delta > 0$ such that

$$|y-z| < \delta \Longrightarrow |f(y)-f(z)| < \varepsilon/2.$$

Choose $M \in \mathbb{Z}^+$ such that $\forall N \ge M$,

$$\frac{2\|f\|_\infty}{(N+1)\sin^2\frac{\delta}{2}} < \varepsilon/2.$$

Since f and K_N are 2π-periodic,

$$\begin{aligned}\sigma_N f(x) &= \int_{-\pi}^{\pi} K_N(x-t)f(t)\,dt && (\tau = x-t)\\ &= \int_{x-\pi}^{x+\pi} K_N(\tau)f(x-\tau)\,d\tau\\ &= \int_{\pi}^{\pi} K_N(\tau)f(x-\tau)\,d\tau.\end{aligned}$$

Then for all $N \geq M$, $\forall x \in [-\pi, \pi]$,

$$\begin{aligned}
|\sigma_N f(x) - f(x)| &= \left| \int_{-\pi}^{\pi} K_N(t) f(x-t)\, dt - \int_{\pi}^{\pi} K_N(t) f(x)\, dt \right| \\
&= \left| \int_{-\pi}^{\pi} K_N(t)[f(x-t) - f(x)]\, dt \right| \\
&\leq \int_{-\pi}^{\pi} |K_N(t)[f(x-t) - f(x)]\, dt| \\
&= \int_{|t|<\delta} K_N(t)|f(x-t) - f(x)|\, dt + \int_{\delta<|t|<\pi} K_N(t)|f(x-t) - f(t)|\, dt \\
&< \frac{\varepsilon}{2} \int_{|t|<\delta} K_N(t)\, dt + 2\|f\|_\infty \int_{\delta<|t|<\pi} \frac{1}{2\pi(N+1)\sin^2\frac{\delta}{2}}\, dt \\
&\leq \frac{\varepsilon}{2} + \frac{2\|f\|_\infty}{(N+1)\sin^2\frac{\delta}{2}} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

□

The same proof can be modified if instead of $K_N(x) \geq 0$, we have $\sup_N \int_{-\pi}^{\pi} |K_N(x)|\, dx < \infty$. But note that

$$\int_{-\pi}^{\pi} |D_N(x)|\, dx \approx \log N.$$

Theorem 16.3. *For all $f \in L^2([-\pi, \pi])$,*

$$\|\sigma_N f\|_2 \leq \|f\|_{L^2}.$$

Proof. Suppose first that $f \in C([-\pi, \pi])$ is 2π-periodic, and extend it to $\mathbb{R}$ using periodicity. Then

$$\begin{aligned}
\sigma_N f(x) &= \int_{-\pi}^{\pi} f(x-t) K_N(t)\, dt \\
\implies \int_{-\pi}^{\pi} |\sigma_N f(x)|^2\, dx &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(x-s)\overline{f(x-t)} K_N(s) K_N(t)\, ds\, dt\, dx \\
&= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} K_N(s) K_N(t) \left[\int_{-\pi}^{\pi} f(x-s)\overline{f(x-t)}\, dx \right] ds\, dt \\
&\leq \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} K_N(s) K_N(t) \|f(-s)\|_{L^2} \|f(-t)\|_{L^2}\, ds\, dt \\
&= \|f\|_{L^2}^2 \int_{-\pi}^{\pi} K_N(s)\, ds \int_{-\pi}^{\pi} K_N(t)\, dt \\
&= \|f\|_{L^2}^2 \\
\implies \|\sigma_N f\| \leq \|f\|_{L_2}.
\end{aligned}$$

Let $f \in L^2([-\pi, \pi])$ An assignment problem in this course says that the set of 2π-periodic continuous functions are dense in $L^2([-\pi, \pi])$. Thus, there exists $\{f_n\}$ 2π-periodic continuous functions such that $\|f_n - f\|_{L^2} \to 0$. Thus, $\|\sigma_N f_n - \sigma_N f\|_{L^2} \to 0$. Thus

$$\|\sigma_N f\|_2 = \lim_{n\to\infty} \|\sigma_N f_n\|_{L^2} \leq \lim_{n\to\infty} \|f_n\|_{L^2} = \|f\|_{L^2}.$$

□

Theorem 16.4. *For all $f \in L^2([-\pi,\pi])$,*

$$\|\sigma_N f - f\|_2 \to 0.$$

In particular, if $\hat{f}(n) = 0$, for all $n \in \mathbb{Z}^+$, then $f = 0$.

Proof. Let $f \in L^2([-\pi,\pi])$. Let $\varepsilon > 0$. There exists $g \in C([-\pi,\pi])$ which is 2π-periodic such that

$$\|f - g\|_2 < \frac{\varepsilon}{3}.$$

□

Since $\sigma_N g \to g$ uniformly on $[-\pi,\pi]$, there exists $M \in \mathbb{Z}^+$ such that $\forall N \geq M$, $\forall x \in [-\pi,\pi]$,

$$|\sigma_N g(x) - g(x)| < \frac{\varepsilon}{3\sqrt{2\pi}}.$$

Then for all $N \geq M$,

$$\begin{aligned}
\|\sigma_N f - f\|_2 &\leq \|\sigma_N(f-g)\|_2 + \|\sigma_N g - g\|_2 + \|g - f\|_2 \qquad (\sigma_N f - \sigma_N g = \sigma_N(f-g)) \\
&\leq 2\|f-g\|_2 + \left(\int_{-\pi}^{\pi} |\sigma_N g(x) - g(x)|^2\, dx\right)^{1/2} \\
&< \frac{2\varepsilon}{3} + \frac{\varepsilon}{3}\left(\int_{-\pi}^{\pi} \frac{1}{2\pi}\, dx\right)^{1/2} \\
&= \varepsilon.
\end{aligned}$$

We've shown the set of exponentials normalized are a maximal orthonormal set, that they are an orthonormal basis of L^2. That is

$$\forall f \in L^2, \quad \|S_N f - f\|_2 \to 0.$$

For a long time it was not necessarily believed that the partial sums converge almost everywhere. However, a deep theorem due to Carleson shows that

$$\forall f \in L^2, \quad S_N f(x) \to f(x), \text{ almost everywhere.}$$

It is said that Carleson spent a decade trying to prove that the negation is true. We have other L^ps spaces, so do the Fourier coefficients converge in other L^p spaces? It is known that

$$\forall 1 < p < \infty, \quad \|S_N f - f\|_p \to 0.$$

When $p = 1, p = \infty$, this is false. For $p = 1$, one can create a counter example where the partial sums diverge almost everywhere. These deeper results require us to do some harmonic analysis. We'll move on now to minimizers over closed convex sets, and their consequences, including the fact that we can identify the dual of a Hilbert space in a canonical way.

17 Lecture 17. Minimizers, Orthogonal Complements, Riesz Representation Theorem

Let's learn how to decompose Hilbert spaces given closed linear subspaces. We'll do this using a similar tool we learned in linear algebra: orthocomplements.

17.1 Length Minimizers

Theorem 17.1. *Suppose $C \subset H$ is a subset of a Hilbert space H such that*

1. *$C \neq \emptyset$,*
2. *C is closed,*
3. *C is convex: if $v_1, v_2 \in C$, and $t \in [0,1]$, then*

$$tv_1 + (1-t)v_2 \in C.$$

Then there exists a unique $v \in C$ such that

$$\|v\| = \inf_{u \in C} \|u\|.$$

Proof. Recall that $a = \inf S$ iff a is a lower bound for S and there exists $\{s_n\} \subset S$, and $s_n \to s$. Let $d = \inf_{u \in C} \|u\|$. Then there exists $\{u_n\} \subset C$ such that

$$\|u_n\| \to d.$$

We claim that $\{u_n\}$ is Cauchy. Let $\varepsilon > 0$. Then there exists $N\ in \mathbb{Z}^+$, $\forall n \geq N$,

$$2\|u_n\|^2 < 2d^2 + \frac{\varepsilon^2}{2}.$$

Then for all $n, m \geq N$, by the parallelogram law,

$$\begin{aligned} \|u_n - u_m\|^2 &= 2\|u_n\|^2 + 2\|u_m\|^2 - 4\left\|\frac{u_n + u_m}{2}\right\|^2 \\ &\leq 2\|u_n\|^2 + 2\|u_m\|^2 - 4d^2 \\ &< 2d^2 + \frac{\varepsilon^2}{2} + 2d^2 + \frac{\varepsilon^2}{2} - 4d^2 = \varepsilon^2. \end{aligned}$$

Then $\exists v \in H$ such that $u_n \to v$. Since C is closed, $v \in C$. Finally,

$$\|v\| = \lim_{n \to \infty} \|u_n\| = d.$$

Thus, $\exists v \in C, \|v\| = d$. Suppose $v, \overline{v} \in C$, and $\|v\| = \|\overline{v}\| = d$. Then by the parallelogram law,

$$\begin{aligned} \|v - \overline{v}\|^2 &= 2\|v\|^2 + 2\|\overline{v}\|^2 - 4\left\|\frac{v + \overline{v}}{2}\right\|^2 \\ &= 4d^2 - 4\left\|\frac{v + \overline{v}}{2}\right\|^2 \\ &\overset{(1)}{\leq} 4d^2 - 4d^2 = 0, \end{aligned}$$

where we argue that $\frac{v+\overline{v}}{2} \in C$ by convexity, and thus it must have greater norm than d. ☐

17.2 Orthocomplements

It's complement, not compliment.

Theorem 17.2. *If H is a Hilbert space, and $W \subset H$ is a subspace, then*

$$W^{\perp} = \{u \in H \mid \langle v, w \rangle = 0, \forall w \in W\}$$

is a closed linear subspace of H. If W is closed, then $H = W \oplus W^{\perp}$

The same direct sum used in linear algebra applies here:

$$\forall u \in H, \exists! w \in W, w^{\perp} \in W^{\perp}, \text{ s.t. } u = w + w^{\perp}.$$

Proof. It's simple to see that $W^{\perp}$ is a subspace of H. Furthermore, $W \cap W^{\perp} = \{0\}$. To show that $W^{\perp}$ is closed, let $\{u_n\}$ be a sequence in $W^{\perp}$, and $u \in H$ such that $u_n \to u$. Let $w \in W$. Then

$$\langle u, w \rangle = \lim_{n \to \infty} \langle u_n, w \rangle = 0.$$

Therefore, $u \in W^{\perp}$. Suppose now that W is closed. if $W = H$, then $W^{\perp} = \{0\}$, and $H = W \oplus \{0\}$ trivially. Assume that $W \neq H$. Let $u \in H \setminus W$. Define $C = u + W$. Not that $u \in C$, so C is nonempty. Also, C is convex, since if $u + w_1 \in C$, $u + w_2 \in C$, and $t \in [0, 1]$, then

$$t(u + w_1) + (1 - t)(u + w_2) = u + \overbrace{(tw_1 + (1 - t)w_2)}^{\in W} \in C.$$

To show C is closed, suppose $u + w_n \to v \in H$. We want to show $v \in C$. If $w + w_n \to v \implies w_n \to v - u$. Since W is closed, $v - u \in W$. Since $v - u \in W$, $v = u + w$, for some $w \in W$, so $v \in C$. Since C is closed and convex, $\exists! v \in C$ such that $\|v\| = \inf_{w \in W} \|u + w\|$. Note that $v \in C$, so $u - v \in W$, and $u = (u - v) + v$. We claim that $v \in W^{\perp}$. Let $w \in W$. Let $f(t) = \|v + tw\|^2 = \|v\|^2 + t^2\|w\|^2 + 2t \operatorname{Re}\langle v, w \rangle$. Then $f(t)$ has a minimum at $t = 0$. Therefore, $f'(0) = 0$, therefore, $\operatorname{Re}\langle v, w \rangle = 0$. Repeat this argument with iw in place of w to get that $\operatorname{Re}\langle v, iw \rangle = \operatorname{Im}\langle v, w \rangle = 0$. Therefore, $\langle v, w \rangle = 0$, so $w \in W^{\perp}$. If $u = w_1 + w_1^{\perp} = w_2 + w_2^{\perp}$, then

$$\overbrace{w_2 - w_1}^{\in W} = \overbrace{w_1^{\perp} - w_2^{\perp}}^{\in W^{\perp}} = 0.$$

□

Theorem 17.3. *If $W \subset H$ is a subspace*

$$\overline{W} = (W^{\perp})^{\perp}.$$

Definition 17.4. A bounded linear operator $P : H \to H$ is a projection if $P^2 = P$.

Theorem 17.5. *Let H be a Hilbert space, and $W \subset H$ be a closed subspace. Then $H = W \oplus W^{\perp}$. The map $\Pi_W : H \to H$ defined by $P_W(v) = w$, where $v = w + w^{\perp}$, is a projection.*

Proof. First we prove that Π_W is linear. If $v_1 = w_1 + w_1^{\perp}, v_2 = w_2 + w_2^{\perp}$, and $\lambda_1, \lambda_2 \in \mathbb{C}$, then

$$\Pi_W(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 w_1 + \lambda_2 v_2 = \lambda_1 \Pi_W(v_1) + \lambda_2 \Pi_W(v_2).$$

Now we must show Π_W is bounded. If $v = w + w^\perp$, then

$$\begin{aligned} \|v\|^2 &= \|w + w^\perp\|^2 \\ &= \|w\|^2 + \|w^\perp\|^2 \\ &\geq \|w\|^2 \\ \implies \|\Pi_W(v)\| &\leq \|v\|. \end{aligned}$$

Finally, if $v = w + w^\perp$, then

$$\Pi_W(\Pi_W(v)) = \Pi_W(w) = w.$$

□

17.3 Riesz Representation Theorem

The only category theory I know is Baire Category theorem, the only representation theory I know is Riesz Representation theorem.

Theorem 17.6. *If H is a Hilbert space, then $\forall f \in H'$, there exists a unique $v \in H$ such that*

$$f(u) = \langle u, v \rangle, \quad u \in H.$$

Proof. We know if v exists, v is unique, since if $f(u) = \langle u, v \rangle = \langle u, v' \rangle$, then

$$\forall u \in H, \langle u, v - v' \rangle = 0, \forall u \in H,$$

so $v - v' = 0$, so $v = v'$. If $f \in H'$ is the identically zero function, then we choose $v = 0$. Suppose $f \neq 0$. Then $\exists u_1 \in H$ such that $f(u_1) \neq 0$. If $u_0 = \frac{u_1}{f(u_1)}$, then $f(u_0) = 1$. Let

$$C = \{u \in H \mid f(u) = 1\}.$$

This set is nonempty, and we can see $C = f^{-1}(\{1\})$. So $\{1\}$ is closed (metric spaces are Hausdorff), and f is bounded and linear, thus continuous, so $f^{-1}(\{1\}) = C$ is closed. Now let's check that C is convex. If $u_1, u_2 \in C$, and $t \in [0, 1]$, then

$$\begin{aligned} f(tu_1 + (1-t)u_2) &= tf(u_1) + (1-t)f(u_2) \\ &= t + 1 - t = 1. \end{aligned}$$

Thus C is convex. Therefore, Theorem 17.1 tells us that $\exists v_0 \in C$ such that

$$\|v_0\| = \inf_{u \in C} \|u\|.$$

Let $v = \frac{v}{\|v_0\|^2}$. Note that $v_0 \neq 0$, since $f(v_0) \neq 0$. We claim that v is the vector satisfying $f(u) = \langle u, v \rangle, \forall u \in H$. Let $N = f^{-1}(\{0\}) = \{w \in H \mid f(w) = 0\}$. Then $C = \{v_0 + w \mid w \in N\}$. Then $\|v_0\| = \inf_{w \in N} \|v_0 + w\|$. By the argument we used above with the function $f(t) = \|v_0 + tw\|^2$, $v_0 \in N^\perp$.

Let $u \in H$. Then

$$f(u - f(u)v_0) = f(u) - f(u)f(v_0) = 0.$$

Thus,

$$u = \underbrace{u - f(u)v_0}_{:=w \in N} + \underbrace{f(u)v_0}_{\in N^\perp}.$$

Thus,

$$\begin{aligned}
\langle u, v \rangle &= \frac{1}{\|v_0\|}\langle u, v_0 \rangle \\
&= \frac{1}{\|v_0\|^2}\left(\langle w, v_0 \rangle + f(u)\langle v_0, v_0 \rangle\right) \\
&= \frac{1}{\|v_0\|^2} f(u)\|v_0\|^2 = f(u).
\end{aligned}$$

□

Soon, we will talk about the analogs of adjoints of linear operators in the infinite dimensional case. We will see that they are not as simple as the finite dimensional case.

18 Lecture 18. The Adjoint of a Bounded Linear Operator on a Hilbert Space

18.1 Adjoints of Bounded Linear Operators

Theorem 18.1. *Let H be a Hilbert space, and $A : H \to H$ be a bounded linear operator. Then there exists a unique bounded linear operator $A^* : H \to H$, which we call the adjoint of A such that $\forall u, v \in H$,*

$$\langle Au, v \rangle = \langle u, A^* v \rangle. \tag{*}$$

Moreover, $\|A^*\| = \|A\|$.

Proof. Uniqueness of A^* follows from (*). Let $v \in H$. Define $f_v : H \to \mathbb{C}$ via

$$f_v(u) = \langle Au, v \rangle.$$

This is a linear map, by the linearity of inner products in the first slots, and linearity of A. If $\|u\| = 1$, then

$$\begin{aligned}
|f_v(u)| = |\langle Au, v \rangle| &\leq \|Au\|\|v\| \\
&\leq \|A\|\|v\|.
\end{aligned}$$

Therefore, $\|f_v\| \leq \|A\|\|v\|$. Thus $f \in H'$. By the Riesz representation theorem, there exists a unique $A^* v \in H$ such that $\forall u \in H$,

$$f_v(u) = \langle u, A^* v \rangle.$$

First, we claim $v \mapsto A^* v$ is linear. Let $v_1, v_2 \in H$, $\lambda_1, \lambda_2 \in \mathbb{C}$. Then $\forall u \in H$,

$$\begin{aligned}
\langle u, A^*(\lambda_1 v_1 + \lambda_2 v_2) \rangle &= \langle Au, \lambda_1 v_1 + \lambda_2 v_2 \rangle \\
&= \overline{\lambda_1}\langle Au, v_1 \rangle + \overline{\lambda_2}\langle Au, v_2 \rangle \\
&= \overline{\lambda_1}\langle u, A^* v_1 \rangle + \overline{\lambda_2}\langle u, A^* v_2 \rangle \\
&= \langle u, \lambda_1 A^* v_1 + \lambda_2 A^* v_2 \rangle.
\end{aligned}$$

This implies that $A^*(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 A^* v_1 + \lambda_2 A^* v_2$. Thus $v \mapsto A^* v$ is linear. Thus this map, which we will call $A^* : H \to H$ is a linear operator. Suppose $\|v\| = 1$. If $A^* v = 0$, then $\|A^* v\| \leq \|A\|$. Suppose $A^* v \neq 0$. Then

$$\begin{aligned} \|A^* v\|^2 &= \langle A^* v, A^* v \rangle \\ &= \langle A A^* v, v \rangle \\ &\leq \|A A^* v\| \|v\| \\ &\leq \|A\| \|A^* v\| \\ \implies \|A^* v\| &\leq \|A\| \\ \implies \|A^*\| &\leq \|A\|. \end{aligned}$$

Note that $\forall u, v \in H$,

$$\begin{aligned} \langle A^* u, v \rangle &= \overline{\langle v, A^* u \rangle} = \overline{\langle Av, u \rangle} = \langle u, Av \rangle \\ \implies \langle u, (A^*)^* v \rangle = \langle A^* u, v \rangle &= \langle u, Av \rangle. \end{aligned}$$

Therefore $(A^*)^* = A$. Thus, $\|A\| = \|(A^*)^*\| \leq \|A^*\|$. Conclude $\|A\| = \|A^*\|$. □

Example 18.2. Suppose $u = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \in \mathbb{C}^n$. Then $(Au)_i = \sum_{j=1}^n A_{ij} u_j$, $A_{ij} \in \mathbb{C}$. Then

$$\begin{aligned} \langle Au, v \rangle &= \sum_{i=1}^n (Au)_i \overline{v_i} \\ &= \sum_{i,j=1}^n A_{ij} u_j \overline{v_i} \\ &= \sum_{j=1}^n u_j \overline{\sum_{i=1}^n \overline{A_{ij}} v_i} \\ &= \sum_{j=1}^n u_j \overline{(A^* v)}, \end{aligned}$$

where $(A^* v)_i = \sum_{j=1}^n \overline{A_{ji}} v_j$. (Recall that $A = (A_{ij}) \implies (A^*)_{ij} = \overline{A_{ji}}$.)

Example 18.3. Suppose $\{A_{ij}\}_{i,j=1}^\infty$ is a double sequence in $\mathbb{C}^n$ such that

$$\sum_{i,j} |A_{ij}|^2 = \lim_{N \to \infty} \sum_{i=1}^N \sum_{j=1}^N |A_{ij}|^2 < \infty.$$

Define $A : \ell^2 \to \ell^2$ by $(A\underline{a})_i = \sum_{j=1}^\infty A_{ij} a_j$, for $a = \{a_j\} \in \ell^2$. Then $A \in \mathcal{B}(\ell^2, \ell^2)$, and $\forall \underline{a}, \underline{b} \in \ell^2$,

$$\langle A\underline{a}, \underline{b} \rangle_{\ell^2} = \sum_i \sum_j A_{ij} a_j \overline{b_i} = \sum_j a_j \left(\overline{\sum_i \overline{A_{ij}} v_i} \right) = \langle \underline{a}, A^* \underline{b} \rangle,$$

where $(A^* \underline{b})_i = \sum_{j=1}^\infty \overline{A_{ji}} b_j$.

Example 18.4. Suppose K is a continuous function in $C([0,1] \times [0,1])$. Define $A : L^2([0,1]) \to L^2([0,1])$ by

$$Af(x) = \int_0^1 K(x,y) f(y)\, dy.$$

Then we can check that

$$A^* g(x) = \int_0^1 \overline{K(y,x)} g(y)\, dy.$$

Theorem 18.5. *Suppose H is a Hilbert space, and $A : H \to H$ is a bounded linear operator. Then*

$$\mathrm{Range}(A)^\perp = \mathrm{Null}(A^*).$$

As a result, suppose that $\mathrm{Range}(A)$ is closed. Then $A : H \to H$ is surjective iff $A^* : H \to H$ is injective.

Proof. We know $v \in \mathrm{Null}(A^*)$ iff $\langle u, A^* v\rangle = 0, \forall u \in H$. This holds iff $\langle Au, v\rangle = 0, \forall u \in H$. That is, v is orthogonal to the range of A. □

18.2 Compactness in a Hilbert Space

At some point, we should have heard of the rank-nullity theorem of linear algebra. It would be great if we could have an equally powerful theorem in the infinite-dimensional setting. To get to something maybe as strong, we need to study compactness in Hilbert spaces.

Definition 18.6. Recall that if X is a metric space, we say $K \subset X$ is compact if every sequence of elements in K admits a convergent subsequence in K.

Or equivalently, every open covering of K admits a finite subcovering.

Example 18.7. Finite subsets of metric spaces are compact, by the pigeonhole principle.

Theorem 18.8 (Heine-Borel)**.** *A subset $K \subset \mathbb{R}^n$ is compact iff it is closed and bounded.*

Example 18.9. The set $[a,b] \subset \mathbb{R}$ is compact for all $a < b, a, b \in \mathbb{R}$.

Heine-Borel does not hold in a general metric space topology.

Example 18.10. Suppose H is a infinite-dimensional Hilbert space. Let $F = \{u \in H \mid\mid \|u\| \le 1\}$. This is a closed and bounded set of H, but it is not compact. Let $\{e_n\}_{n=1}^\infty$ be a countable orthonormal subset of H. Then $\forall n \neq k$, $\|e_n - e_k\|^2 = \|e_n\|^2 + \|e_k\|^2 + 2\,\mathrm{Re}\langle e_n, e_k\rangle = 2$. Intuitively we can see that e_n cannot admit a convergent subsequence, since convergent subsequences need to be Cauchy.

We know in the metric space topology, a compact subset is closed. What additional condition can we check that a subset of a Hilbert space is compact? Actually, we have seen this question posed in analysis. Recall the Arzela-Ascoli theorem.

Theorem 18.11 (Arzela-Ascoli)**.** *Let (X, d) be a metric space, and let $E \subset X$ be compact. Let $f_n \in C(E)$, $\forall n \in \mathbb{Z}^+$. Suppose $\{f_n\}$ is pointwise bounded and equicontinuous on E. Then $\{f_n\}$ is uniformly bounded on E and $\{f_n\}$ contains a uniformly convergent subsequence.*

We will relate this idea of equicontinuity to the Hilbert space sense.

Definition 18.12. Let H be a Hilbert space. A subset K of H has equismall tails with respect to a countable orthonormal subset $\{e_k\} \subset H$ if $\forall \varepsilon > 0$, $\exists N \in \mathbb{Z}^+$ such that $\forall v \in K$,

$$\sum_{k>N} |\langle v, e_k \rangle|^2 < \varepsilon^2.$$

Example 18.13. If K is $\{v_1, v_2, \dots, v_n\} \implies K$ has equismall tails with respect to any orthonormal subset $\{e_k\}$.

Theorem 18.14. *Let H be a Hilbert space, and let $\{v_n\}$ be a sequence with $v_n \to v$. Let $\{e_k\}$ be a countable orthonormal subset.*

1. *The set $K = \{v_n \mid n \in \mathbb{N}\} \cup \{v\}$ is compact.*
2. *The set K has equismall tails with respect to $\{e_k\}$.*

Proof. Actually, Prof. Casey leaves part 1 as an exercise. It's an analysis exercise, which I sort of remember, so let's prove it. When K is finite, this is obvious. If K is countably infinite, let $\{k_n\}$ be a sequence in K. If $\{k_n\}$ takes on only finitely many values, then of course it admits a convergent subsequence.

Suppose that $\{k_n\}$ takes on countably infinitely many values. Relabel the entries in $\{k_n\}$ with the sequential values of $\{v_n\}$ as follows. Let $v_{n_1} = k_1$. If there does not exist an index n_2, with $n_2 > n_1$ such that $v_{n_2} = k_\ell$, for some $\ell > n_1$, then $\{k_n\}$ contains finitely many values. Choose $v_{n_2} = k_\ell$, where $\ell > n_1$. Continuing in this fashion to the n_kth index, if there does not exist an index n_{k+1}, with $n_{k+1} > n_k$, and $v_{n_{k+1}} = k_\ell$, for some $\ell > n_k$, then conclude that $\{k_n\}$ contains finitely many values. Then by induction we create a subsequence $\{v_{n_k}\}$ which is a subsequence of $\{k_n\}$ as well as $\{v_n\}$. Since $v_n \to v$, $v_{n_k} \to v$, so conclude that K is compact.

Now that we've stretched our brains with analysis, let's continue with the proof of part 2. Let $\varepsilon > 0$. Since $v_n \to v$, $\exists M \in \mathbb{N}$ such that $\forall n \geq M$, $\|v_n - v\| < \varepsilon/2$. Choose $N \in \mathbb{N}$ so large that

$$\sum_{k>N} |\langle v, e_k \rangle|^2 + \max_{1 \leq n \leq M-1} \sum_{k \geq N} |\langle v_n, e_k \rangle|^2 < \frac{\varepsilon^2}{4}.$$

There are only finitely many terms here, so we can choose our N large enough so that it makes the $n = 1$ term smaller that $\varepsilon^2/8$, the $n = 2$ term smaller that $\varepsilon^2/16$, and so on. Then

$$\sum_{k>N} |\langle v, e_k \rangle|^2 < \varepsilon^2/4 < \varepsilon^2,$$

and for all $1 \leq n \leq M - 1$,

$$\sum_{k>N} |\langle v_n, e_k \rangle|^2 < \varepsilon^2/4 < \varepsilon^2.$$

If $n \geq M$, by Bessel's inequality,

$$\left(\sum_{k>N} |\langle v_n, e_k \rangle|^2 \right)^{1/2} = \left(\sum_{k>N} |\langle v_n - v, e_k \rangle + \langle v, e_k \rangle|^2 \right)^{1/2}, \tag{1}$$

and this is the ℓ^2 norm of the sum of two sequences indexed by k, so by the triangle inequality, this is bounded by

$$\left(\sum_{k>N} |\langle v_n - v, e_k \rangle|^2 \right)^{1/2} + \left(\sum_{k>N} |\langle v, e_k \rangle|^2 \right)^{1/2}.$$

The second term is at most $\varepsilon/2$, and the first term is bounded by Bessel's inequality by $|v_n - v|$. Since we chose N large enough so that the norm is less than $\varepsilon/2$, we indeed have that this is bounded by ε. ☐

Next time, we'll show that if a subset of a separable Hilbert space is closed, bounded, and has equismall tails with respect to an orthonormal basis, which we know exists, then we have compactness, and we'll rephrase this in a way that doesn't involve Hilbert spaces, and go from there.

19 Lecture 19. Compact Subsets of a Hilbert Space and Finite-Rank Operators

Last time, we proved that in a Hilbert space H, for a sequence $\{v_n\}$ in H such that $v_n \to v$, the set K of all of the values it takes on, unioned with v is compact, and that the set K has equismall tails with respect to any countable orthonormal subset of H. That is $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $\forall \tilde{v} \in K$,

$$\sum_{k>N} |\langle \tilde{v}, e_k \rangle|^2 < \varepsilon^2.$$

We now are equipped to prove the following theorem about compact subsets of Hilbert spaces.

Theorem 19.1. *Let H be a separable Hilbert space, and let $\{e_k\}$ be an orthonormal basis of H. Then $K \subset H$ is compact if and only if K is closed, bounded, and has equismall tails with respect to $\{e_k\}$.*

Proof. Suppose first that K is compact. Then K is closed and bounded by metric space theory. Suppose by way of contradiction, K does not have equismall tails with respect to $\{e_k\}$. That is, $\exists \varepsilon_0 > 0$ such that $\forall N \in \mathbb{N}$, $\exists u_N \in K$ such that

$$\sum_{k>N} |\langle u_N, e_k \rangle|^2 \geq \varepsilon_0^2. \tag{*}$$

Note that $\{u_N\}_N$ is a sequence in K. Therefore, there exists a subsequence $\{v_n\}_n$ of $\{u_N\}_N$ which is convergent to a point in $v \in K$. Then $\forall n \in \mathbb{N}$,

$$\sum_{k>n} |\langle v_n, e_k \rangle|^2 \geq \varepsilon_0^2.$$

Then $\{v_n \mid n \in \mathbb{N}\} \cup \{v\}$ does not have equismall tails, which contradicts Theorem 18.14. Thus K has equismall tails with respect to $\{e_k\}$.

For the converse, assume that K is closed, bounded, and has equismall tails with respect to $\{e_k\}$. Let $\{u_n\}_n$ be a sequence in K. Since K is closed, if $\{u_n\}$ converges, we are done. By Bolzano-Weierstrauss, which holds in the complex case, a bounded sequence in $\mathbb{C}$ has a convergent subsequence. Since K is bounded, $\exists C \geq 0$ such that

$$\forall n \in \mathbb{N}, \|u_n\| \leq C.$$

Therefore, $\forall k, \forall n, |\langle u_n, e_k \rangle| \leq \|u_n\|\|e_k\| \leq C$. That is, $\forall k \in \mathbb{N}$, $\{\langle u_n, e_k \rangle\}_n$ is a bounded sequence in $\mathbb{C}$. Since $\{\langle u_n, e_1 \rangle\}_n$ is bounded, by Bolzano-Weierstrauss, there exists a subsequence

$$\{\langle u_{n_1(j)}, e_1 \rangle\}_j$$

of $\{\langle u_n, e_1\rangle\}_n$ which converges in $\mathbb{C}$. Since $\{\langle u_{n_1(j)}, e_2\rangle\}_j$ is a bounded sequence, there exists a subsequence

$$\{\langle u_{n_2(j)}, e_2\rangle\}_j$$

of $\{\langle u_{n_1(j)}, e_2\rangle\}$ which converges. Note that

$$\lim_{j\to\infty} \langle u_{n_2(j)}, e_1\rangle \text{ exists,}$$
$$\lim_{j\to\infty} \langle u_{n_2(j)}, e_2\rangle \text{ exists.}$$

Then $\forall \ell$, there exists a subsequence $\{n_\ell(j)\}_j$ of $\{n_{\ell-1}(j)\}_j$ such that $\forall 1 \le k \le \ell$,

$$\lim_{j\to\infty} \langle u_{n_\ell}(j), e_k\rangle_j \text{ exists.}$$

Pick

$$v_\ell = u_{n_\ell(\ell)}, \quad \ell = 1, 2, 3, \ldots.$$

(The diagonal along this sequence.) Then $\{v_\ell\}_\ell$ is a subsequence of $\{u_n\}_n$ such that $\forall k$,

$$\{\langle v_\ell, e_k\rangle\}_\ell \text{ converges.}$$

We claim that $\{v_\ell\}_\ell$ is Cauchy, thus convergent, since H is Hilbert. Let $\varepsilon > 0$. Since K has equismall tails, $\exists N \in \mathbb{N}$ such that $\forall \ell \in \mathbb{N}$,

$$\sum_{k>N} |\langle v_\ell, e_k\rangle|^2 < \frac{\varepsilon^2}{16}.$$

Since the N sequences $\{\langle v_\ell, e_1\rangle\}_\ell, \ldots, \{\langle v_\ell, e_N\rangle\}_\ell$ converge, $\exists M \in \mathbb{N}$, such that $\forall \ell, m \ge M$,

$$\sum_{k=1}^{N} |\langle v_\ell, e_k\rangle - \langle v_m, e_k\rangle|^2 < \frac{\varepsilon^2}{4}.$$

Then $\forall \ell, m \ge M$,

$$\|v_\ell - v_m\| = \left(\sum_{k=1}^{N} |\langle v_\ell - v_m, e_k\rangle|^2 + \sum_{k>N} |\langle v_\ell - v_m, e_k\rangle|^2\right)^{1/2} \qquad \text{(Bessel's inequality)}$$
$$\le \left(\sum_{k=1}^{N} |\langle v_\ell - v_m, e_k\rangle|^2\right)^{1/2} + \left(\sum_{k>N} |\langle v_\ell - v_m, e_k\rangle|^2\right)^{1/2}. \qquad (2)$$

By our choice of M, the first term is bounded above by $\varepsilon/2$, and we can use the ℓ^2 triangle inequality to bound the second term:

$$< \frac{\varepsilon}{2} + \left(\sum_{k>N} |\langle v_\ell, e_k\rangle|^2\right)^{1/2} + \left(\sum_{k>N} |\langle v_m, e_k\rangle|^2\right)^{1/2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon.$$

Thus our subsequence is Cauchy, thus convergent, this K is compact. □

Note. We didn't use anything about $\{e_k\}$ being a basis, so if we in fact have a compact set of H, it has equismall tails with respect to any orthonormal subset.