MATH455 - Analysis 4

Abstract Metric, Topological Spaces; Functional Analysis.

Based on lectures from Winter 2025 by Prof. Jessica Lin.
Notes by Louis Meunier

Contents

§1 Abstract Metric and Topological Spaces

§1.1 Review of Metric Spaces

Throughout fix X a nonempty set.

↪**Definition 1.1** (Metric): $\rho : X \times X \to \mathbb{R}$ is called a *metric*, and thus (X, ρ) a *metric space*, if for all $x, y, z \in X$,
- $\rho(x, y) \geq 0$,
- $\rho(x, y) = 0 \Leftrightarrow x = y$,
- $\rho(x, y) = \rho(y, x)$, and
- $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$.

↪**Definition 1.2** (Norm): Let X a linear space. A function $\| \cdot \| : X \to [0, \infty)$ is called a *norm* if for all $u, v \in X$ and $\alpha \in \mathbb{R}$,
- $\|u\| = 0 \Leftrightarrow u = 0$,
- $\|u + v\| \leq \|u\| + \|v\|$, and
- $\|\alpha u\| = |\alpha| \; \|u\|$.

Remark 1.1: A norm induces a metric by $\rho(x, y) := \|x - y\|$.

↪**Definition 1.3**: Given two metrics ρ, σ on X, we say they are *equivalent* if $\exists\, C > 0$ such that $\frac{1}{C}\sigma(x, y) \leq \rho(x, y) \leq C\sigma(x, y)$ for every $x, y \in X$. A similar definition follows for equivalence of norms.

Given a metric space (X, ρ), then, we have the notion of
- open balls $B(x, r) = \{y \in X : \rho(x, y) < r\}$,
- open sets (subsets of X with the property that for every $x \in X$, there is a constant $r > 0$ such that $B(x, r) \subseteq X$), closed sets, closures, and
- *convergence*.

↪**Definition 1.4** (Convergence): $\{x_n\} \subseteq X$ converges to $x \in X$ if $\lim_{n \to \infty} \rho(x_n, x) = 0$.

We have several (equivalent) notions, then, of continuity; via sequences, $\varepsilon - \delta$ definition, and by pullbacks (inverse images of open sets are open).

↪**Definition 1.5** (Uniform Continuity): $f : (X, \rho) \to (Y, \sigma)$ uniformly continuous if f has a "modulus of continuity", i.e. there is a continuous function $\omega : [0, \infty) \to [0, \infty)$ such that

$$\sigma(f(x_1), f(x_2)) \le \omega(\rho(x_1, x_2))$$

for every $x_1, x_2 \in X$.

Remark 1.2: For instance, we say f Lipschitz continuous if there is a constant $C > 0$ such that $\omega(\cdot) = C(\cdot)$. Let $\alpha \in (0, 1)$. We say f α-Holder continuous if $\omega(\cdot) = C(\cdot)^\alpha$ for some constant C.

↪**Definition 1.6** (Completeness): We say (X, ρ) *complete* if every Cauchy sequence in (X, ρ) converges to a point in X.

Remark 1.3: If (X, ρ) complete and $E \subseteq X$, then (E, ρ) is complete iff E closed in X.

§1.2 Compactness, Separability

↪**Definition 1.7** (Open Cover, Compactness): $\{X_\lambda\}_{\lambda \in \Lambda} \subseteq 2^X$, where X_λ open in X and Λ an arbitrary index set, an *open cover* of X if for every $x \in X$, $\exists\, \lambda \in \Lambda$ such that $x \in X_\lambda$.

X is *compact* if every open cover of X admits a compact subcover. We say $E \subseteq X$ compact if (E, ρ) compact.

↪**Definition 1.8** (Totally Bounded, ε-nets): (X, ρ) *totally bounded* if $\forall\, \varepsilon > 0$, there is a finite cover of X of balls of radius ε. If $E \subseteq X$, an ε-*net* of E is a collection $\{B(x_i, \varepsilon)\}_{i=1}^N$ such that $E \subseteq \bigcup_{i=1}^N B(x_i, \varepsilon)$ and $x_i \in X$ (note that x_i need not be in E).

↪**Definition 1.9** (Sequentially Compact): (X, ρ) *sequentially compact* if every sequence in X has a convergent subsequence whose limit is in X.

↪**Definition 1.10** (Relatively/Pre- Compact): $E \subseteq X$ *relatively compact* if $\overline{E}$ compact.

↪**Theorem 1.1**: TFAE:
1. X complete and totally bounded;
2. X compact;
3. X sequentially compact.

Remark 1.4: $E \subseteq X$ relatively compact if every sequence in E has a convergent subsequence.

Let $f : (X, \rho) \to (Y, \sigma)$ continuous with (X, ρ) compact. Then,

- $f(X)$ compact in Y;
- if $Y = \mathbb{R}$, the max and min of f over X are achieved;
- f is uniformly continuous.

Let $C(X) := \{f : X \to \mathbb{R} \mid f \text{ continuous}\}$ and $\|f\|_\infty := \max_{x \in X}|f(x)|$ the sup (max, in this case) norm. Then,

↪**Theorem 1.2**: Let (X, ρ) compact. Then, $(C(X), \|\cdot\|_\infty)$ is complete.

Proof. Let $\{f_n\} \subseteq C(X)$ Cauchy with respect to $\|\cdot\|_\infty$. Then, there exists a subsequence $\left\{f_{n_k}\right\}$ such that for each $k \geq 1$, $\|f_{n_{k+1}} - f_{n_k}\|_\infty \leq 2^{-k}$ (to construct this subsequence, let $n_1 \geq 1$ be such that $\|f_n - f_{n_1}\|_\infty < \frac{1}{2}$ for all $n \geq n_1$, which exists since $\{f_n\}$ Cauchy. Then, for each $k \geq 1$, define inductively n_{k+1} such that $n_{k+1} > n_k$ and $\|f_n - f_{n_{k+1}}\|_\infty < \frac{1}{2^{k+1}}$ for each $n \geq n_{k+1}$. Then, for any $k \geq 1$, $\|f_{n_{k+1}} - f_{n_k}\|_\infty < 2^{-k}$, since $n_{k+1} > n_k$.).

Let $j \in \mathbb{N}$. Then, for any $k \geq 1$,

$$\|f_{n_{k+j}} - f_{n_k}\|_\infty \leq \sum_{\ell=k}^{k+j-1} \|f_{n_{\ell+1}} - f_{n_\ell}\|_\infty \leq \sum_{\ell} 2^{-\ell}$$

and hence for each $x \in X$, with $c_k := f_{n_k}(x)$,

$$|c_{k+j} - c_k| \leq \sum_{\ell=k}^{\infty} 2^{-\ell}.$$

The RHS is the tail of a converging series, and thus $|c_{k+j} - c_k| \to 0$ as $k \to \infty$ i.e. $\{c_k\}$ a Cauchy sequence, in $\mathbb{R}$. $(\mathbb{R}, |\cdot|)$ complete, so $\lim_{k\to\infty} c_k =: f(x)$ exists for each $x \in X$. So, for each $x \in X$, we find

$$|f_{n_k}(x) - f(x)| \leq \sum_{\ell=k}^{\infty} 2^{-\ell},$$

and since the RHS is independent of x, we may pass to the sup norm, and find

$$\|f_{n_k} - f\|_\infty \leq \sum_{\ell=k}^{\infty} 2^{-\ell},$$

with the RHS $\to 0$ as $k \to \infty$. Hence, $f_{n_k} \to f$ in $C(X)$ as $k \to \infty$. In other words, we have uniform convergence of $\left\{f_{n_k}\right\}$. Each $\left\{f_{n_k}\right\}$ continuous, and thus f also continuous, and thus $f \in C(X)$.

It remains to show convergence along the whole sequence. Suppose otherwise. Then, there is some $\alpha > 0$ and a subsequence $\left\{f_{n_j}\right\} \subseteq \{f_n\}$ such that $\|f_{n_j} - f\|_\infty >$

$\alpha > 0$ for every $j \geq 1$. Then, let k be sufficiently large such that $\|f - f_{n_k}\|_\infty \leq \frac{\alpha}{2}$. Then, for every $j \geq 1$ and k sufficiently large,

$$\begin{aligned}\|f_{n_j} - f_{n_k}\|_\infty &\geq \|f_{n_j} - f\|_\infty - \|f - f_{n_k}\|_\infty \\ &> \alpha - \frac{\alpha}{2} > 0,\end{aligned}$$

which contradicts the Cauchy-ness of $\{f_n\}$, completing the proof. ■

↪**Definition 1.11** (Density/Separability): A set $D \subseteq X$ is called *dense* in X if for every nonempty open subset $A \subseteq X$, $D \cap A \neq \varnothing$. We say X *separable* if there is a countable dense subset of X.

Remark 1.5: If A dense in X, then $\overline{A} = X$.

↪**Proposition 1.1**: If X compact, X separable.

Proof. Since X compact, it is totally bounded. So, for $n \in \mathbb{N}$, there is some K_n and $\{x_i\} \subseteq X$ such that $X \subseteq \bigcup_{i=1}^{K_n} B\left(x_i, \frac{1}{n}\right)$. Then, $D = \bigcup_{n=1}^{\infty} \bigcup_{i=1}^{K_n} \{x_i\}$ countable and dense in X. ■

§1.3 Arzelà-Ascoli

The goal in this section is to find conditions for a sequence of functions $\{f_n\} \subseteq C(X)$ to be precompact, namely, to have a uniformly convergent subsequence.

↪**Corollary 1.1**: Any Cauchy sequence converges if it has a convergent subsequence.

Proof. Let $\{x_n\}$ be a Cauchy sequence in a metric space (X, ρ) with convergent subsequence $\left\{x_{n_k}\right\}$ which converges to some $x \in X$. Fix $\varepsilon > 0$. Let $N \geq 1$ be such that if $m, n \geq N$, $\rho(x_n, x_m) < \frac{\varepsilon}{2}$. Let $K \geq 1$ be such that if $k \geq K$, $\rho\left(x_{n_k}, x\right) < \frac{\varepsilon}{2}$. Let $n, n_k \geq \max\{N, K\}$, then

$$\rho(x, x_n) \leq \rho\left(x, x_{n_k}\right) + \rho\left(x_{n_k}, x_n\right) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

■

↪**Definition 1.12** (Equicontinuous): A family $\mathcal{F} \subseteq C(X)$ is called *equicontinuous* at $x \in X$ if $\forall\, \varepsilon > 0$ there exists a $\delta = \delta(x, \varepsilon) > 0$ such that if $\rho(x, x') < \delta$ then $|f(x) - f(x')| < \varepsilon$ for every $f \in \mathcal{F}$.

Remark 1.6: $\mathcal{F}$ equicontinuous at x iff every $f \in \mathcal{F}$ share the same modulus of continuity.

↪**Definition** **1.13** (Pointwise/uniformly bounded): $\{f_n\}$ pointwise bounded if $\forall\, x \in X$, $\exists\, M(x) > 0$ such that $|f_n(x)| \leq M(x)\,\forall\, n$, and uniformly bounded if such an M exists independent of x.

↪**Lemma** **1.1** (Arzelà-Ascoli Lemma): Let X separable and let $\{f_n\} \subseteq C(X)$ be pointwise bounded and equicontinuous. Then, there is a function f and a subsequence $\left\{f_{n_k}\right\}$ which converges pointwise to f on all of X.

Proof. Let $D = \left\{x_j\right\}_{j=1}^{\infty} \subseteq X$ be a countable dense subset of X. Since $\{f_n\}$ p.w. bounded, $\{f_n(x_1)\}$ as a sequence of real numbers is bounded and so by the Bolzano-Weierstrass (BW) Theorem there is a convergent subsequence $\left\{f_{n(1,k)}(x_1)\right\}_k$ that converges to some $a_1 \in \mathbb{R}$. Consider now $\left\{f_{n(1,k)}(x_2)\right\}_k$, which is again a bounded sequence of $\mathbb{R}$ and so has a convergent subsequence, call it $\left\{f_{n(2,k)}(x_2)\right\}_k$ which converges to some $a_2 \in \mathbb{R}$. Note that $\left\{f_{n(2,k)}\right\} \subseteq \left\{f_{n(1,k)}\right\}$, so also $f_{n(2,k)}(x_1) \to a_1$ as $k \to \infty$. We can repeat this procedure, producing a sequence of real numbers $\{a_\ell\}$, and for each $j \in \mathbb{N}$ a subsequence $\left\{f_{n(j,k)}\right\}_k \subseteq \{f_n\}$ such that $f_{n(j,k)}(x_\ell) \to a_\ell$ for each $1 \leq \ell \leq j$. Define then

$$f : D \to \mathbb{R}, f(x_j) := a_j.$$

Consider now

$$f_{n_k} := f_{n(k,k)}, k \geq 1,$$

the "diagonal sequence", and remark that $f_{n_k}(x_j) \to a_j = f(x_j)$ as $k \to \infty$ for every $j \geq 1$. Hence, $\left\{f_{n_k}\right\}_k$ converges to f on D, pointwise.

We claim now that $\left\{f_{n_k}\right\}$ converges on all of X to some function $f : X \to \mathbb{R}$, pointwise. Put $g_k := f_{n_k}$ for notational convenience. Fix $x_0 \in X$, $\varepsilon > 0$, and let $\delta > 0$ be such that if $x \in X$ such that $\rho(x, x_0) < \delta$, $|g_k(x) - g_k(x_0)| < \frac{\varepsilon}{3}$ for every $k \geq 1$, which exists by equicontinuity. Since D dense in X, there is some $x_j \in D$ such that $\rho(x_j, x_0) < \delta$. Then, since $g_k(x_j) \to f(x_j)$ (pointwise), $\{g_k(x_j)\}_k$ is Cauchy and so there is some $K \geq 1$ such that for every $k, \ell \geq K$, $|g_\ell(x_j) - g_k(x_j)| < \frac{\varepsilon}{3}$. And hence, for every $k, \ell \geq K$,

$$|g_k(x_0) - g_\ell(x_0)| \leq |g_k(x_0) - g_k(x_j)| + |g_k(x_j) - g_\ell(x_j)| + |g_\ell(x_j) - g_\ell(x_0)| < \varepsilon,$$

so namely $\{g_k(x_0)\}_k$ Cauchy as a sequence in $\mathbb{R}$. Since $\mathbb{R}$ complete, then $\{g_k(x_0)\}_k$ also converges, to, say, $f(x_0) \in \mathbb{R}$. Since x_0 was arbitrary, this means there is some function $f : X \to \mathbb{R}$ such that $g_k \to f$ pointwise on X as we aimed to show. ■

↪**Definition 1.14** (Uniformly Equicontinuous): $\mathcal{F} \subseteq C(X)$ is said to be uniformly equicontinuous if for every $\varepsilon < 0$, there exists a $\delta > 0$ such that $\forall\, x, y \in X$ with $\rho(x, y) < \delta$, $|f(x) - f(y)| < \varepsilon$ for every $f \in \mathcal{F}$. That is, every function in $\mathcal{F}$ has the same modulus of continuity.

↪**Proposition 1.2** (Sufficient Conditions for Uniform Equicontinuity):
1. $\mathcal{F} \subseteq C(X)$ uniformly Lipschitz
2. $\mathcal{F} \subseteq C(X) \cap C^1(X)$ has a uniform L^∞ bound on the first derivative
3. $\mathcal{F} \subseteq C(X)$ uniformly Holder continuous
4. (X, ρ) compact and $\mathcal{F}$ equicontinuous

PROOF.

1. If $C > 0$ is such that $|f(x) - f(y)| \leq C\rho(x, y)$ for every $x, y \in X$ and $f \in \mathcal{F}$, then for $\varepsilon > 0$, let $\delta = \frac{\varepsilon}{C}$, then if $\rho(x, y) \leq \delta$, $|f(x) - f(y)| \leq C\delta < \varepsilon$, and δ independent of x (and f) since it only depends on C which is independent of x, y, f, etc.
3. Akin to 1.

■

↪**Theorem 1.3** (Arzelà-Ascoli): Let (X, ρ) a compact metric space and $\{f_n\} \subseteq C(X)$ be a uniformly bounded and (uniformly) equicontinuous family of functions. Then, $\{f_n\}$ is pre-compact in $C(X)$, i.e. there exists $\left\{f_{n_k}\right\} \subseteq \{f_n\}$ such that f_{n_k} is uniformly convergent on X.

PROOF. Since (X, ρ) compact it is separable and so by the lemma there is a subsequence $\left\{f_{n_k}\right\}$ that converges pointwise on X. Denote by $g_k := f_{n_k}$ for notational convenience.

We claim $\{g_k\}$ uniformly Cauchy. Let $\varepsilon > 0$. By uniform equicontinuity, there is a $\delta > 0$ such that $\rho(x, y) < \delta \Rightarrow |g_k(x) - g_k(y)| < \frac{\varepsilon}{3}$. Since X compact it is totally bounded so there exists $\{x_i\}_{i=1}^N$ such that $X \subseteq \bigcup_{i=1}^N B(x_i, \delta)$. For every $1 \leq i \leq N$, $\{g_k(x_i)\}$ converges by the lemma hence is Cauchy in $\mathbb{R}$. So, there exists a K_i such that for every $k, \ell \geq K_i$ $|g_k(x_i) - g_\ell(x_i)| \leq \frac{\varepsilon}{3}$. Let $K := \max\{K_i\}$. Then for every $\ell, k \leq K$, $|g_k(x_i) - g_\ell(x_i)| \leq \frac{\varepsilon}{3}$ for every $i = 1, ..., N$. So, for all $x \in X$, there is some x_i such that $\rho(x, x_i) < \delta$, and so for every $k, \ell \geq K$,

$$\begin{aligned} |g_k(x) - g_\ell(x)| \leq |g_k(x) - g_k(x_i)| \\ + |g_k(x_i) - g_\ell(x_i)| \\ + |g_\ell(x_i) - g_\ell(x)| < \varepsilon, \end{aligned}$$

the first and last follow by the equicontinuity and the second from the lemma. This holds for every x and thus $\|g_k - g_\ell\|_\infty < \varepsilon$, so $\{g_k\}$ Cauchy in $C(X)$. But $C(X)$ complete so converges in the space. ■

Remark 1.7: If $K \subseteq X$ a compact set, then K bounded and closed.

↪**Theorem 1.4**: Let (X, ρ) compact and $\mathcal{F} \subseteq C(X)$. Then, $\mathcal{F}$ a compact subspace of $C(X)$ iff $\mathcal{F}$ closed, uniformly bounded, and (uniformly) equicontinuous.

Proof. ($\Leftarrow$) Let $\{f_n\} \subseteq \mathcal{F}$. By Arzelà-Ascoli Theorem, there exists a subsequence $\{f_{n_k}\}$ that converges uniformly to some $f \in C(X)$. Since $\mathcal{F}$ closed, $f \in \mathcal{F}$ and so $\mathcal{F}$ sequentially compact hence compact.

($\Rightarrow$) $\mathcal{F}$ compact so closed and bounded in $C(X)$. To prove equicontinuous, we argue by contradiction. Suppose otherwise, that $\mathcal{F}$ not-equicontinuous at some $x \in X$. Then, there is some $\varepsilon_0 > 0$ and $\{f_n\} \subseteq \mathcal{F}$ and $\{x_n\} \subseteq X$ such that $|f_n(x_n) - f_n(x)| \geq \varepsilon_0$ while $\rho(x, x_n) < \frac{1}{n}$. Since $\{f_n\}$ bounded and $\mathcal{F}$ compact, there is a subsequence $\{f_{n_k}\}$ that converges to f uniformly. Let K be such that $\forall\, k \geq K$, $\|f_{n_k} - f\|_\infty \leq \frac{\varepsilon_0}{3}$. Then,

$$\begin{aligned} |f(x_{n_k}) - f\,| &\geq \big|\, |f(x_{n_k}) - f_{n_k}(x_{n_k})| - |f_{n_k}(x_{n_k}) - f_{n_k}(x)| - |f_{n_k}(x) - f(x)|\,\big| \\ &\geq \frac{\varepsilon_0}{3}, \end{aligned}$$

while $\rho(x_{n_k}, x) \leq \frac{1}{n_k}$, so f cannot be continuous at x, a contradiction. ■

§1.4 Baire Category Theorem

↪**Definition 1.15** (Hollow/Nowhere Dense): We say a set $E \subseteq X$ *hollow* if $\operatorname{int}(E) = \varnothing$. We say a set $E \subseteq X$ *nowhere dense* if its closure is hollow, i.e. $\operatorname{int}(\overline{E}) = \varnothing$.

Remark 1.8: Notice that E hollow $\Leftrightarrow$ E^c dense, since $\operatorname{int}(E) = \varnothing \Rightarrow (\operatorname{int}(E))^c = \overline{E^c} = X$.

↪**Theorem 1.5** (Baire Category Theorem): Let X be a complete metric space.

(a) Let $\{F_n\}$ a collection of closed hollow sets. Then, $\bigcup_{n=1}^\infty F_n$ also hollow.

(b) Let $\{\mathcal{O}_n\}$ a collection of open dense sets. Then, $\bigcap_{n=1}^\infty \mathcal{O}_n$ also dense.

Proof. Notice that $(a) \Leftrightarrow (b)$ by taking complements. We prove (b).

Put $G := \bigcap_{n=1}^\infty \mathcal{O}_n$. Fix $x \in X$ and $r > 0$, then to show density of G is to show $G \cap B(x, r) \neq \varnothing$.

Since $\mathcal{O}_1$ dense, then $\mathcal{O}_1 \cap B(x, r)$ nonempty and in particular open. So, let $x_1 \in X$ and $r_1 < \frac{1}{2}$ such that $\overline{B}(x, r_1) \subseteq B(x, 2r_1) \subseteq \mathcal{O}_1 \cap B(x, r)$.

Similarly, since $\mathcal{O}_2$ dense, $\mathcal{O}_2 \cap B(x_1, r_1)$ open and nonempty so there exists $x_2 \in X$ and $r_2 < 2^{-2}$ such that $\overline{B}(x_2, r_2) \subseteq \mathcal{O}_2 \cap B(x_1, r_1)$.

Repeat in this manner to find $x_n \in X$ with $r_n < 2^{-n}$ such that $\overline{B}(x_n, r_n) \subseteq \mathcal{O}_n \cap B(x_{n-1}, r_{n-1})$ for any $n \in \mathbb{N}$. This creates a sequence of sets

$$\overline{B}(x_1, r_1) \supseteq \overline{B}(x_2, r_2) \supseteq \cdots,$$

with $r_n \to 0$. Hence, the sequence of points $\{x_n\}$ Cauchy and since X complete, $x_j \to x_0 \in X$, so in particular

$$\{x_0\} = \bigcap_{n=1}^{\infty} \overline{B}(x_n, r_n),$$

hence $x_0 \in \mathcal{O}_n$ for every n and thus $G \cap B(x, r)$ nonempty. ■

↪**Corollary 1.2**: Let X complete and $\{F_n\}$ a sequence of closed sets in X. If $X = \bigcup_{n \geq 1} F_n$, there is some n_0 such that $\operatorname{int}\left(F_{n_0}\right) \neq \varnothing$.

Proof. If not, violates BCT since X is not hollow in itself; $\operatorname{int}(X) = X$. ■

↪**Corollary 1.3**: Let X complete and $\{F_n\}$ a sequence of closed sets in X. Then, $\bigcup_{n=1}^{\infty} \partial F_n$ hollow.

Proof. We claim $\operatorname{int}(\partial F_n) = \varnothing$. Suppose not, then there exists some $B(x_0, r) \subseteq \partial F_n$. Then $x_0 \in \partial F_n$ but $B(x_0, r) \cap F_n^c = \varnothing$, a contradiction. So, since ∂F_n closed and $\partial F_n \cap B(x_0, r) = \varnothing$ for every such ball, by BCT $\bigcup_{n=1}^{\infty} \partial F_n$ must be hollow. ■

1.4.1 Applications of Baire Category Theorem

↪**Theorem 1.6**: Let $\mathcal{F} \subset C(X)$ where X complete. Suppose $\mathcal{F}$ pointwise bounded. Then, there exists a nonempty, open set $\mathcal{O} \subseteq X$ such that $\mathcal{F}$ uniformly bounded on $\mathcal{O}$.

Proof. Let

$$\begin{aligned} E_n &:= \{x \in X : |f(x)| \leq n \; \forall \, f \in \mathcal{F}\} \\ &= \bigcap_{f \in \mathcal{F}} \underbrace{\{x : |f(x)| \leq n\}}_{\text{closed}}. \end{aligned}$$

Since $\mathcal{F}$ pointwise bounded, for every $x \in X$ there is some $M_x > 0$ such that $|f(x)| \leq M_x$ for every $f \in \mathcal{F}$. Hence, for every $n \in \mathbb{N}$ such that $n \geq M_x$, $x \in E_n$ and thus $X = \bigcup_{n=1}^{\infty} E_n$.

E_n closed and hence by the previous corollaries there is some n_0 such that $\operatorname{int}\left(E_{n_0}\right) \neq \varnothing$ and hence there is some $r > 0$ and $x_0 \in X$ such that $B(x_0, r) \subseteq E_{n_0}$. Then, for every $x \in B(x_0, r)$, $|f(x)| \leq n_0$ for every $f \in \mathcal{F}$, which gives our desired non-empty open set upon which $\mathcal{F}$ uniformly bounded. ■

↪**Theorem 1.7**: Let X complete, and $\{f_n\} \subseteq C(X)$ such that $f_n \to f$ pointwise on X. Then, there exists a dense subset $D \subseteq X$ such that $\{f_n\}$ equicontinuous on D and f continuous on D.

PROOF. For $m, n \in \mathbb{N}$, let

$$\begin{aligned} E(m,n) &:= \left\{x \in X : |f_j(x) - f_k(x)| \leq \frac{1}{m} \,\forall\, j,k \geq n\right\} \\ &= \bigcap_{j,k \geq n} \left\{x : |f_j(x) - f_k(x)| \leq \frac{1}{m}\right\}. \end{aligned}$$

The union of the boundaries of these sets are hollow, hence $D := \left(\bigcup_{m,n\geq 1} \partial E(m,n)\right)^c$ is dense. Then, if $x \in D \cap E(m,n)$, then $x \in (\partial E(m,n))^c$ implies $x \in \mathrm{int}(E(m,n))$.

We claim $\{f_n\}$ equicontinuous on D. Let $x_0 \in D$ and $\varepsilon > 0$. Let $\frac{1}{m} \leq \frac{\varepsilon}{4}$. Then, since $\{f_n(x_0)\}$ convergent it is therefore Cauchy (in $\mathbb{R}$). Hence, there is some N such that $|f_j(x_0) - f_k(x_0)| \leq \frac{1}{m}$ for every $j, k \geq N$, so $x_0 \in D \cap E(m,N)$ hence $x_0 \in \mathrm{int}(E(m,N))$.

Let $B(x_0, r) \subseteq E(m,N)$. Since f_N continuous at x_0 there is some $\delta > 0$ such that $\delta < r$ and

$$|f_N(x) - f_N(x_0)| < \frac{1}{m} \,\forall\, x \in B(x_0, \delta),$$

and hence

$$\begin{aligned} |f_j(x) - f_j(x_0)| &\leq |f_j(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f_j(x_0)| \\ &\leq \frac{3}{m} \leq \frac{3}{4}\varepsilon, \end{aligned}$$

for every $x \in B(x_0, \delta)$ and $j \geq N$, where the first, last bounds come from Cauchy and the middle from continuity of f_N. Hence, we've show $\{f_n\}$ equicontinuous at x_0 since δ was independent of f.

In particular, this also gives for every $x \in B(x_0, \delta)$ the limit

$$\frac{3}{4}\varepsilon > \lim_{j\to\infty} |f_j(x) - f_j(x_0)| = |f(x) - f(x_0)|,$$

so f continuous on D. ■

§1.5 Topological Spaces

Throughout, assume $X \neq \varnothing$.

↪**Definition 1.16** (Topology): Let $X \neq \varnothing$. A *topology* $\mathcal{T}$ on X is a collection of subsets of X, called *open sets*, such that

- $X, \varnothing \in \mathcal{T}$;
- If $\{E_n\} \subseteq \mathcal{T}$, $\bigcap_{n=1}^{N} E_n \in \mathcal{T}$ (closed under *finite* intersections);
- If $\{E_n\} \subseteq \mathcal{T}$, $\bigcup_n E_n \in \mathcal{T}$ (closed under *arbitrary* unions).

If $x \in X$, a set $E \in \mathcal{T}$ containing x is called a neighborhood of x.

↪**Proposition 1.3**: $E \subseteq X$ open $\Leftrightarrow$ for every $x \in E$, there is a neighborhood of x contained in E.

Proof. $\Rightarrow$ is trivial by taking the neighborhood to be E itself. $\Leftarrow$ follows from the fact that, if for each x we let $\mathcal{U}_x$ a neighborhood of x contained in E, then

$$E = \bigcup_{x \in E} \mathcal{U}_x,$$

so E open being a union of open sets. ■

⊛ **Example 1.1**: Every metric space induces a natural topology given by open sets under the metric. The *discrete topology* is given by $\mathcal{T} = 2^X$ (and is actually induced by the discrete metric), and is the largest topology. The *trivial topology* $\{\varnothing, X\}$ is the smallest. The *relative topology* defined on a subset $Y \subseteq X$ is given by $\mathcal{T}_Y := \{E \cap Y : E \in \mathcal{T}\}$.

↪**Definition 1.17** (Base): Given a topological space $(X, \mathcal{T})$, let $x \in X$. A collection $\mathcal{B}_x$ of neighborhoods of x is called a *base* of $\mathcal{T}$ at x if for every neighborhood $\mathcal{U}$ of x, there is a set $B \in \mathcal{B}_x$ such that $B \subseteq \mathcal{U}$.

We say a collection $\mathcal{B}$ a base for all of $\mathcal{T}$ if for every $x \in X$, there is a base for x, $\mathcal{B}_x \subseteq \mathcal{B}$.

↪**Proposition 1.4**: If $(X, \mathcal{T})$ a topological space, then $\mathcal{B} \subseteq \mathcal{T}$ a base for $\mathcal{T}$ $\Leftrightarrow$ every nonempty open set $\mathcal{U} \in \mathcal{T}$ can be written as a union of elements of $\mathcal{B}$.

Proof. $\Rightarrow$ If $\mathcal{U}$ open, then for $x \in \mathcal{U}$ there is some basis element B_x contained in $\mathcal{U}$. So in particular $\mathcal{U} = \bigcup_{x \in \mathcal{U}} B_x$.

$\Leftarrow$ Let $x \in \mathcal{U}$ and $\mathcal{B}_x := \{B \in \mathcal{B} \mid x \in B\}$. Then, for every neighborhood of x, there is some B in $\mathcal{B}_x$ such that $B \subseteq \mathcal{U}$ so $\mathcal{B}_x$ a base for $\mathcal{T}$ at x. ■

Remark 1.9: A base $\mathcal{B}$ defines a unique topology, $\{\varnothing, \cup \mathcal{B}_x\}$.

↪**Proposition 1.5**: $\mathcal{B} \subseteq 2^X$ a base for a topology on $X \Leftrightarrow$
- $X = \bigcup_{B \in \mathcal{B}} B$
- If $B_1, B_2 \in \mathcal{B}$ and $x \in B_1 \cap B_2$, then there is a $B \in \mathcal{B}$ such that $x \in B \subseteq B_1 \cap B_2$.

PROOF. ($\Rightarrow$) If $\mathcal{B}$ a base, then X open so $X = \cup_B B$. If $B_1, B_2 \in \mathcal{B}$, then $B_1 \cap B_2$ open so there must exist some $B \subseteq B_1 \cap B_2$ in $\mathcal{B}$.

($\Leftarrow$) Let

$$\mathcal{T} = \{\mathcal{U} \mid \forall x \in \mathcal{U}, \exists B \in \mathcal{B} \text{ with } x \in B \subseteq \mathcal{U}\}.$$

One can show this a topology on X with $\mathcal{B}$ as a base. ■

↪**Definition 1.18**: If $\mathcal{T}_1 \subsetneq \mathcal{T}_2$, we say $\mathcal{T}_1$ *weaker/coarser* and $\mathcal{T}_2$ *stronger/finer*.

Given a subset $S \subseteq 2^X$, define

$$\mathcal{T}(S) = \bigcap \text{all topologies containing } S = \text{unique weakest topology containing } S$$

to be the topology *generated* by S.

↪**Proposition 1.6**: If $S \subseteq 2^X$,

$$\mathcal{T}(S) = \bigcup \{\text{finite intersections of elts of } S\}.$$

We call S a "subbase" for $\mathcal{T}(S)$ (namely, we allow finite intersections of elements in S to serve as a base for $\mathcal{T}(S)$).

PROOF. Let $\mathcal{B} := \{X, \text{finite intersections of elements of } S\}$. We claim this a base for $\mathcal{T}(S)$. ■

↪**Definition 1.19** (Point of closure/accumulation point): If $E \subseteq X, x \in X$, x is called a *point of closure* if $\forall \mathcal{U}_x, \mathcal{U}_x \cap E \neq \varnothing$. The collection of all such sets is called the *closure* of E, denoted $\overline{E}$. We say E *closed* if $E = \overline{E}$.

↪**Proposition 1.7**: Let $E \subseteq X$, then
- $\overline{E}$ closed,
- $\overline{E}$ is the smallest closed set containing E,
- E open $\Leftrightarrow E^c$ closed.

§1.6 Separation, Countability, Separability

↪**Definition 1.20**: A neighborhood of a set $K \subseteq X$ is any open set containing K.

↪**Definition 1.21** (Notions of Separation): We say $(X, \mathcal{T})$:

- *Tychonoff Separable* if $\forall\, x, y \in X, \exists\, \mathcal{U}_x, \mathcal{U}_y$ such that $y \notin \mathcal{U}_x, x \notin \mathcal{U}_y$
- *Hausdorff Separable* if $\forall\, x, y \in X$ can be separated by two disjoint open sets i.e. $\exists\, \mathcal{U}_x \cap \mathcal{U}_y = \varnothing$
- *Normal* if Tychonoff and in addition any 2 disjoint closed sets can be separated by disjoint neighborhoods.

Remark 1.10: Metric space $\subseteq$ normal space $\subseteq$ Hausdorff space $\subseteq$ Tychonoff space.

↪**Proposition 1.8**: Tychonoff $\Leftrightarrow\ \forall\, x \in X$, $\{x\}$ closed.

Proof. For every $x \in X$,

$$\begin{aligned}\{x\} \text{ closed} &\Leftrightarrow \{x\}^c \text{ open}\\ &\Leftrightarrow \forall\, y \in \{x\}^c, \exists\, \mathcal{U}_y \subseteq \{x\}^c\\ &\Leftrightarrow \forall\, y \neq x, \exists\, \mathcal{U}_y \text{ s.t. } x \notin \mathcal{U}_y,\end{aligned}$$

and since this holds for every x, X Tychonoff. ■

↪**Proposition 1.9**: Every metric space normal.

Proof. Define, for $F \subseteq X$, the function

$$\operatorname{dist}(F, x) := \inf\{\rho(x, x') \mid x' \in F\}.$$

Notice that if F closed and $x \notin F$, then $\operatorname{dist}(F, x) > 0$ (since F^c open so there exists some $B(x, \varepsilon) \subseteq F^c$ so $\rho(x, x') \geq \varepsilon$ for every $x' \in F$). Let F_1, F_2 be closed disjoint sets, and define

$$\begin{aligned}\mathcal{O}_1 &:= \{x \in X \mid \operatorname{dist}(F_1, x) < \operatorname{dist}(F_2, x)\},\\ \mathcal{O}_2 &:= \{x \in X \mid \operatorname{dist}(F_1, x) > \operatorname{dist}(F_2, x)\}.\end{aligned}$$

Then, $F_1 \subseteq \mathcal{O}_1, F_2 \subseteq \mathcal{O}_2$, and $\mathcal{O}_1 \cap \mathcal{O}_2 = \varnothing$. If we show $\mathcal{O}_1, \mathcal{O}_2$ open, we'll be done.

Let $x \in \mathcal{O}_1$ and $\varepsilon > 0$ such that $\operatorname{dist}(F_1, x) + \varepsilon \leq \operatorname{dist}(F_2, x)$. I claim that $B\left(x, \frac{\varepsilon}{5}\right) \subseteq \mathcal{O}_1$. Let $y \in B\left(x, \frac{\varepsilon}{5}\right)$. Then,

$$
\begin{aligned}
\operatorname{dist}(F_2, y) &\geq \rho(y, y') - \frac{\varepsilon}{5} \qquad \text{for some } y' \in F_2 \\
&\geq \rho(x, y') - \rho(x, y) + \frac{\varepsilon}{5} \qquad \text{reverse triangle inequality} \\
&\geq \operatorname{dist}(F_2, x) - \frac{2\varepsilon}{5} \\
&\geq \operatorname{dist}(F_1, x) + \varepsilon - \frac{2\varepsilon}{5} \\
&\geq \rho(x, \tilde{y}) + \frac{2\varepsilon}{5} \qquad \text{for some } \tilde{y} \in F_1 \\
&\geq \rho(y, \tilde{y}) - \rho(y, x) + \frac{2\varepsilon}{5} \qquad \text{reverse triangle inequality} \\
&\geq \rho(y, \tilde{y}) - \frac{\varepsilon}{5} + \frac{2\varepsilon}{5} \\
&\geq \operatorname{dist}(F_1, y) + \frac{\varepsilon}{5} > \operatorname{dist}(F_1, y),
\end{aligned}
$$

hence, $y \in \mathcal{O}_1$ and thus $\mathcal{O}_1$ open. Similar proof follows for $\mathcal{O}_2$. ■

↪**Proposition 1.10**: Let X Tychonoff. Then X normal $\Leftrightarrow$ $\forall\, F \subseteq X$ closed and neighborhood $\mathcal{U}$ of F, there exists an open set $\mathcal{O}$ such that

$$F \subseteq \mathcal{O} \subseteq \overline{\mathcal{O}} \subseteq \mathcal{U}.$$

This is called the "nested neighborhood property" of normal spaces.

Proof. ($\Rightarrow$) Let F closed and $\mathcal{U}$ a neighborhood of F. Then, F and $\mathcal{U}^c$ closed disjoint sets so by normality there exists $\mathcal{O}, \mathcal{V}$ disjoint open neighborhoods of $F, \mathcal{U}^c$ respectively. So, $\mathcal{O} \subseteq \mathcal{V}^c$ hence $\overline{\mathcal{O}} \subseteq \overline{\mathcal{V}}^c = \mathcal{V}^c$ and thus

$$F \subseteq \mathcal{O} \subseteq \overline{\mathcal{O}} \subseteq \mathcal{V}^c \subseteq \mathcal{U}.$$

($\Leftarrow$) Let A, B be disjoint closed sets. Then, B^c open and moreover $A \subseteq B^c$. Hence, there exists some open set $\mathcal{O}$ such that $A \subseteq \mathcal{O} \subseteq \overline{\mathcal{O}} \subseteq B^c$, and thus $B \subseteq \overline{\mathcal{O}}^c$. Then, $\mathcal{O}$ and $\overline{\mathcal{O}}^c$ are disjoint open neighborhoods of A, B respectively so X normal. ■

↪**Definition 1.22** (Separable): A space X is called *separable* if it contains a countable dense subset.

↪**Definition 1.23** (1st, 2nd Countable): A topological space $(X, \mathcal{T})$ is called

- *1st countable* if there is a countable base at each point; and
- *2nd countable* if there is a countable base for all of $\mathcal{T}$.

⊛ **Example 1.2**: Every metric space is first countable; for $x \in X$ let $\mathcal{B}_x = \{B(x, \frac{1}{n}) \mid n \in \mathbb{N}\}$.

↪**Proposition 1.11**: Every 2nd countable space is separable.

↪**Definition 1.24** (Convergence): Let $\{x_n\} \subseteq X$. Then, we say $x_n \to x$ in $\mathcal{T}$ if for every neighborhood $\mathcal{U}_x$, there exists an N such that $\forall\, n \geq N$, $x_n \in \mathcal{U}_x$.

Remark 1.11: In general spaces, such a limit may not be unique. For instance, under the trivial topology, the only nonempty neighborhood is the whole space, so every sequence converges to every point in the space.

↪**Proposition 1.12**: Let $(X, \mathcal{T})$ be Hausdorff. Then, all limits are unique.

Proof. Suppose otherwise, that $x_n \to$ both x and y. If $x \neq y$, then since X Hausdorff there are disjoint neighborhoods $\mathcal{U}_x, \mathcal{U}_y$ containing x, y. But then x_n cannot be on both $\mathcal{U}_x$ and $\mathcal{U}_y$ for sufficiently large n, contradiction. ■

↪**Proposition 1.13**: Let X be 1st countable and $E \subseteq X$. Then, $x \in \overline{E} \Leftrightarrow$ there exists $\{x_j\} \subseteq E$ such that $x_j \to x$.

Proof. ($\Rightarrow$) Let $\mathcal{B}_x = \{B_j\}$ be a base for X at $x \in \overline{E}$. Wlog, $B_j \supseteq B_{j+1}$ for every $j \geq 1$ (by replacing with intersections, etc if necessary). Hence, $B_j \cap E \neq \varnothing$ for every j. Let $x_j \in B_j \cap E$, then by the nesting property $x_j \to x$ in $\mathcal{T}$.

($\Leftarrow$) Suppose otherwise, that $x \notin \overline{E}$. Let $\{x_j\} \in E_j$. Then, $\overline{E}^c$ open, and contains x. Then, $\overline{E}^c$ a neighborhood of x but does not contain any x_j so $x_j \nrightarrow x$. ■

§1.7 Continuity and Compactness

↪**Definition 1.25**: Let $(X, \mathcal{T}), (Y, \mathcal{S})$ be two topological spaces. Then, a function $f : X \to Y$ is said to be continuous at x_0 if for every neighborhood $\mathcal{O}$ of $f(x_0)$ there exists a neighborhood $\mathcal{U}(x_0)$ such that $f(\mathcal{U}) \subseteq \mathcal{O}$. We say f continuous on X if it is continuous at every point in X.

↪**Proposition 1.14**: f continuous $\Leftrightarrow$ $\forall\, \mathcal{O}$ open in Y, $f^{-1}(\mathcal{O})$ open in X.

↪**Definition 1.26** (Weak Topology): Consider $\mathcal{F} := \{f_\lambda : X \to X_\lambda\}_{\lambda \in \Lambda}$ where X, X_λ topological spaces. Then, let

$$S := \{f_\lambda^{-1}(\mathcal{O}_\lambda) \mid f_\lambda \in \mathcal{F}, \mathcal{O}_\lambda \in X_\lambda\} \subseteq X.$$

We say that the topology $\mathcal{T}(S)$ generated by S is the *weak topology* for X induced by the family $\mathcal{F}$.

↪**Proposition 1.15**: The weak topology is the weakest topology in which each f_λ continuous on X.

⊛ **Example 1.3**: The key example of the weak topology is given by the product topology. Consider $\{X_\lambda\}_{\lambda \in \Lambda}$ a collection of topological spaces. We can defined a "natural" topology on the product $X := \prod_{\lambda \in \Lambda} X_\lambda$ by consider the weak topology induced by the family of projection maps, namely, if $\pi_\lambda : X \to X_\lambda$ a coordinate-wise projection and $\mathcal{F} = \{\pi_\lambda : \lambda \in \Lambda\}$, then we say the weak topology induced by $\mathcal{F}$ is the *product topology* on X. In particular, a base for this topology is given, by previous discussions,

$$\mathcal{B} = \left\{\bigcap_{j=1}^{n} \pi_{\lambda_j}^{-1}(\mathcal{O}_j)\right\} = \left\{\prod_{\lambda \in \Lambda} \mathcal{U}_\lambda : \mathcal{U}_\lambda \text{ open and all by finitely many } U_\lambda\text{'s} = X_\lambda\right\}.$$

↪**Definition 1.27** (Compactness): A space X is said to be *compact* if every open cover of X admits a finite subcover.

↪**Proposition 1.16**:
- Closed subsets of compact spaces are compact
- X compact $\Leftrightarrow$ if $\{F_k\} \subseteq X$-nested and closed, $\cap_{k=1}^{\infty} F_k \neq \varnothing$.
- Continuous images of compact sets are compact
- Continuous real-valued functions on a compact topological space achieve their min, max.

↪**Proposition 1.17**: Let K compact be contained in a Hausdorff space X. Then, K closed in X.

Proof. We show K^c open. Let $y \in K^c$. Then for every $x \in K$, there exists disjoint open sets $\mathcal{U}_{xy}, \mathcal{O}_{xy}$ containing y, x respectively. Then, it follows that $\{\mathcal{O}_{xy}\}_{x \in K}$ an open cover of K, and since K compact there must exist some finite subcover, $K \subseteq \bigcup_{i=1}^{N} \mathcal{O}_{x_i y}$. Let $E := \bigcap_{i=1}^{N} \mathcal{U}_{x_i y}$. Then, E is an open neighborhood of y with $E \cap \mathcal{O}_{x_i y} = \varnothing$ for every

$i = 1, \ldots N$. Thus, $E \subseteq \bigcap_{i=1}^{N} \mathcal{O}_{x_i y}^c = \left(\bigcup_{i=1}^{N} \mathcal{O}_{x_i y}\right)^c \subseteq K^c$ so since y was arbitrary K^c open. ■

↪**Definition 1.28** (Sequential Compactness): We say $(X, \mathcal{T})$ *sequentially compact* if every sequence in X has a converging subsequence with limit contained in X.

↪**Proposition 1.18**: Let $(X, \mathcal{T})$ second countable. Then, X compact $\Leftrightarrow$ sequentially compact.

Proof. ($\Rightarrow$) Let $\{x_k\} \subseteq X$ and put $F_n := \overline{\{x_k \mid k \geq n\}}$. Then, $\{F_n\}$ defines a sequence of closed and nested subsets of X and, since X compact, $\bigcap_{n=1}^{\infty} F_n$ nonempty. Let x_0 in this intersection. Since X 2nd and so in particular 1st countable, let $\{B_j\}$ a (wlog nested) countable base at x_0. $x_0 \in F_n$ for every $n \geq 1$ so each B_j must intersect some F_n. Let n_j be an index such that $x_{n_j} \in B_j$. Then, if $\mathcal{U}$ a neighborhood of x_0, there exists some N such that $B_j \subseteq \mathcal{U}$ for every $j \geq N$ and thus $\left\{x_{n_j}\right\} \subseteq B_N \subseteq \mathcal{U}$, so $x_{n_j} \to x_0$ in X.

($\Leftarrow$) Remark that since X second countable, every open cover of X certainly has a countable subcover by intersecting a given cover with our countable basis. So, assume we have a countable cover $X \subseteq \bigcup_{n=1}^{\infty} \mathcal{O}_n$ and suppose towards a contradiction that no finite subcover exists. Then, for every $n \geq 1$, there exists some $m(n) \geq n$ such that $\mathcal{O}_{m(n)} \setminus \bigcup_{i=1}^{n} \mathcal{O}_i \neq \varnothing$. Let x_n in this set for every $n \geq 1$. Since X sequentially compact, there exists a convergent subsequence $\left\{x_{n_k}\right\} \subseteq \{x_n\}$ such that $x_{n_k} \to x_0$ in X, so there exists some $\mathcal{O}_N$ such that $x_0 \in \mathcal{O}_N$. But by construction, $x_{n_k} \notin \mathcal{O}_N$ if $n_k \geq N$, and we have a contradiction. ■

↪**Theorem 1.8**: If X compact and Hausdorff, X normal.

Proof. We show that any closed set F and any point $x \notin F$ can be separated by disjoint open sets. Then, the proof in the more general case follows.

For each $y \in X$, X is Hausdorff so there exists disjoint open neighborhoods $\mathcal{O}_{xy}$ and $\mathcal{U}_{xy}$ of x, y respectively. Then, $\left\{\mathcal{U}_{xy} \mid y \in F\right\}$ defines an open cover of F. Since F closed and thus, being a subset of a compact space, compact, there exists a finite subcover $F \subseteq \bigcup_{i=1}^{N} \mathcal{U}_{xy_i}$. Put $\mathcal{N} := \bigcap_{i=1}^{N} \mathcal{O}_{xy_i}$. This is an open set containing x, with $\mathcal{N} \cap \bigcup_{i=1}^{N} \mathcal{U}_{xy_i} = \varnothing$ hence F and x separated by $\mathcal{N}, \bigcup_{i=1}^{N} \mathcal{U}_{xy_i}$. ■

§1.8 Connected Topological Spaces

↪**Definition 1.29** (Separate): 2 non-empty sets $\mathcal{O}_1, \mathcal{O}_2$ *separate* X if $\mathcal{O}_1$, $\mathcal{O}_2$ disjoint and $X = \mathcal{O}_1 \cup \mathcal{O}_2$.

↪**Definition 1.30** (Connected): We say X *connected* if it cannot be separated.

Remark 1.12: Note that if X can be separated, then $\mathcal{O}_1, \mathcal{O}_2$ are closed as well as open, being complements of each other.

↪**Proposition 1.19**: Let $f : X \to Y$ continuous. Then, if X connected, so is $f(X)$.

PROOF. Suppose otherwise, that $f(X) = \mathcal{O}_1 \sqcup \mathcal{O}_2$ for nonempty, open, disjoint $\mathcal{O}_1, \mathcal{O}_2$. Then, $X = f^{-1}(\mathcal{O}_1) \sqcup f^{-1}(\mathcal{O}_2)$, and each of these inverse images remain nonempty and open in X, so this a contradiction to the connectedness of X. ■

Remark 1.13: On $\mathbb{R}$, $C \subseteq \mathbb{R}$ connected $\Leftrightarrow$ an interval $\Leftrightarrow$ convex.

↪**Definition 1.31** (Intermediate Value Property): We say X has the intermediate value property (IVP) if $\forall\, f \in C(X)$, $f(X)$ an interval.

↪**Proposition 1.20**: X has IVP $\Leftrightarrow$ X connected.

PROOF. ($\Leftarrow$) If X connected, $f(X)$ connected in $\mathbb{R}$ hence an interval.

($\Rightarrow$) Suppose otherwise, that $X = \mathcal{O}_1 \sqcup \mathcal{O}_2$. Then define the function $f : X \to \mathbb{R}$ by $x \mapsto \begin{cases} 1 \text{ if } x \in \mathcal{O}_2 \\ 0 \text{ if } x \in \mathcal{O}_1 \end{cases}$. Then, for every $A \subseteq \mathbb{R}$,

$$f^{-1}(A) = \begin{cases} \varnothing & \text{if } \{0,1\} \not\subseteq A \\ \mathcal{O}_1 & \text{if } 0 \in A \\ \mathcal{O}_2 & \text{if } 1 \in A \\ X & \text{if } \{0,1\} \subseteq A \end{cases},$$

which are all open sets, hence f continuous. But $f(X) = \{0,1\}$ which is not an interval, hence the IVP fails and so X must be connected. ■

↪**Definition 1.32** (Arcwise/Path Connected): X *arc connected/path connected* if $\forall\, x, y \in X$, there exists a continuous function $f : [0,1] \to X$ such that $f(0) = x$, $f(1) = y$.

↪**Proposition 1.21**: Arc connected $\Rightarrow$ connected.

PROOF. Suppose otherwise, $X = \mathcal{O}_1 \sqcup \mathcal{O}_2$. Let $x \in \mathcal{O}_1, y \in \mathcal{O}_2$ and define a continuous function $f : [0,1] \to X$ such that $f(0) = x$ and $f(1) = y$. Then, $f^{-1}(\mathcal{O}_i)$ each open, nonempty and disjoint for $i = 1, 2$, but

$$f^{-1}(\mathcal{O}_1) \sqcup f^{-1}(\mathcal{O}_2) = [0,1],$$

a contradiction to the connectedness of $[0,1]$. ■

§1.9 Urysohn's Lemma and Urysohn's Metrization Theorem

We present the main lemma of this section first, but need more tools before proving it.

↪**Lemma 1.2** (Urysohn's): Let $A, B \subseteq X$ closed and disjoint subsets of a normal space X. Then, $\forall\, [a, b] \subseteq \mathbb{R}$, there exists a continuous function $f : [a, b] \to \mathbb{R}$ such that $f(X) \subseteq [a, b]$, $f|_A = a$ and $f|_B = b$.

Remark 1.14: We have a partial converse of this statement as well:

↪**Proposition 1.22**: Let X Tychonoff and suppose X satisfies the properties of Urysohn's Lemma. Then, X normal.

PROOF. Let A, B be closed nonempty disjoint subsets. Let $f : X \to \mathbb{R}$ continuous such that $f|_A = 0$, $f|_B = 1$ and $0 \leq f \leq 1$. Let I_1, I_2 be two disjoint open intervals in $\mathbb{R}$ with $0 \in I_1$ and $1 \in I_2$. Then, $f^{-1}(I_1)$ open and contains A, and $f^{-1}(I_2)$ open and contains B. Moreover, $f^{-1}(I_1) \cap f^{-1}(I_2) = \varnothing$; hence, $f^{-1}(I_1), f^{-1}(I_2)$ disjoint open neighborhoods of A, B respectively, so indeed X normal. ■

↪**Definition 1.33** (Normally Ascending): Let $(X, \mathcal{T})$ a topological space and $\Lambda \subseteq \mathbb{R}$. A collection of open sets $\{\mathcal{O}_\lambda\}_{\lambda \in \Lambda}$ is said to be *normally ascending* if $\forall\, \lambda_1, \lambda_2 \in \Lambda$,

$$\overline{\mathcal{O}_{\lambda_1}} \subseteq \mathcal{O}_{\lambda_2} \text{ if } \lambda_1 < \lambda_2.$$

↪**Lemma 1.3**: Let $\Lambda \subseteq (a, b)$ a dense subset, and let $\{\mathcal{O}_\lambda\}_{\lambda \in \Lambda}$ a normally ascending collection of subsets of X. Let $f : X \to \mathbb{R}$ defined such that

$$f(x) = \begin{cases} b & \text{if } x \in \left(\bigcup_{\lambda \in \Lambda} \mathcal{O}_\lambda\right)^c \\ \inf\{\lambda \in \Lambda \mid x \in \mathcal{O}_\lambda\} & \text{else} \end{cases}.$$

Then, f continuous.

PROOF. We claim $f^{-1}(-\infty, c)$ and $f^{-1}(c, \infty)$ open for every $c \in \mathbb{R}$. Since such sets define a subbase for $\mathbb{R}$, it suffices to prove continuity on these sets. We show just the first for convenience. Notice that since $f(x) \in [a, b]$, if $c \in (a, b)$ then $f^{-1}(-\infty, c) = f^{-1}[a, c)$, so really it suffices to show that $f^{-1}[a, c)$ open to complete the proof.

Suppose $x \in f^{-1}([a, c])$ so $a \leq f(x) < c$. Let $\lambda \in \Lambda$ be such that $a < \lambda < f(x)$. Then, $x \notin \mathcal{O}_\lambda$. Let also $\lambda' \in \Lambda$ such that $f(x) < \lambda' < c$. By density of Λ, there exists a $\varepsilon > 0$ such that $f(x) + \varepsilon \in \Lambda$, so in particular

$$\overline{\mathcal{O}}_{f(x)+\varepsilon} \subseteq \mathcal{O}_{\lambda'} \Rightarrow x \in \mathcal{O}_{\lambda'},$$

by nesting. So, repeating this procedure, we find

$$f^{-1}([a,c)) \subseteq \bigcup_{a \leq \lambda < \lambda' < c} \mathcal{O}_{\lambda'} \setminus \overline{\mathcal{O}}_{\lambda},$$

noticing the set on the right is open. By similar reasoning, the opposite inclusion holds and we have equality. Hence, f continuous. ■

↪**Lemma 1.4**: Let X normal, $F \subseteq X$ closed, and $\mathcal{U}$ a neighborhood of F. Then, for any $(a,b) \subseteq \mathbb{R}$, there exists a dense subset $\Lambda \subseteq (a,b)$ and a normally ascending collection $\{\mathcal{O}_\lambda\}_{\lambda \in \Lambda}$ such that

$$F \subseteq \mathcal{O}_\lambda \subseteq \overline{\mathcal{O}}_\lambda \subseteq \mathcal{U}, \qquad \forall\, \lambda \in \Lambda.$$

Remark 1.15: This is essentially a generalization of the nested neighborhood property, and indeed the proof essentially just uses this property repeatedly to construct the collection $\{\mathcal{O}_\lambda\}$.

Proof. Without loss of generality, we assume $(a,b) = (0,1)$, for the two intervals are homeomorphic, i.e. the function $f : (0,1) \to \mathbb{R}$, $f(x) := a(1-x) + bx$ is continuous, invertible with continuous inverse and with $f(0) = a$, $f(1) = b$ so a homeomorphism.

Let

$$\Lambda := \left\{\frac{m}{2^n} \mid m,n \in \mathbb{N} \mid 1 \leq m \leq 2^{n-1}\right\} = \bigcup_{n \in \mathbb{N}} \underbrace{\left\{\frac{m}{2^n} \mid m \in \mathbb{N}, 1 \leq m \leq 2^{n-1}\right\}}_{=:\Lambda_n},$$

which is clearly dense in $(0,1)$. We need now to define our normally ascending collection. We do so by defining on each Λ_1 and proceding inductively.

For Λ_1, since X normal, let $\mathcal{O}_{1/2}$ be such that $F \subseteq \mathcal{O}_{1/2} \subseteq \overline{\mathcal{O}}_{1/2} \subseteq \mathcal{U}$, which exists by the nested neighborhood property.

For $\Lambda_2 = \{\frac{1}{4}, \frac{3}{4}\}$, we use the nested neighborhood property again, but first with F as the closed set and $\mathcal{O}_{1/2}$ an open neighborhood of it, and then with $\overline{\mathcal{O}}_{1/2}$ as the closed set and $\mathcal{U}$ an open neighborhood of it. In this way, we find

$$\underbrace{F \subseteq \mathcal{O}_{1/4} \subseteq \overline{\mathcal{O}}_{1/4} \subseteq \mathcal{O}_{1/2}}_{\text{nested nbhd}} \subseteq \overbrace{\overline{\mathcal{O}}_{1/2} \subseteq \mathcal{O}_{3/4} \subseteq \overline{\mathcal{O}}_{3/4} \subseteq \mathcal{U}}^{\text{nested nbhd}}.$$

We repeat in this manner over all of Λ, in the end defining a normally ascending collection $\{\mathcal{O}_\lambda\}_{\lambda \in \Lambda}$. ■

Proof *(Of Urysohn's Lemma, Lem. 1.2)*. Let $F = A$ and $\mathcal{U} = B^c$ as in the previous lemma Lem. 1.4. Then, there is some dense subset $\Lambda \subseteq (a,b)$ and a normally ascending collection $\{\mathcal{O}_\lambda\}_{\lambda \in \Lambda}$ such that $A \subseteq \mathcal{O}_\lambda \subseteq \overline{\mathcal{O}}_\lambda \subseteq B^c$ for every $\lambda \in \Lambda$. Let $f(x)$ as in the previous previous lemma, Lem. 1.3. Then, if $x \in B$, $B \subseteq \left(\bigcup_{\lambda \in \Lambda} \mathcal{O}_\lambda\right)^c$ and so $f(x) = b$.

Otherwise if $x \in A$, then $x \in \bigcap_{\lambda \in \Lambda} \mathcal{O}_\lambda$ and thus $f(x) = \inf\{\lambda \in \Lambda\} = a$. By the first lemma, f continuous, so we are done. ■

↪**Theorem 1.9** (Urysohn's Metrization Theorem): Let X be a second countable topological space. Then, X is metrizable (that is, there exists a metric on X that induces the topology) if and only if X normal.

PROOF. (⇒) We have already showed, every metric space is normal.

(⇐) Let $\{\mathcal{U}_n\}$ be a countable basis for $\mathcal{T}$ and put

$$A := \left\{(n, m) \in \mathbb{N} \times \mathbb{N} \mid \overline{\mathcal{U}}_n \subseteq \mathcal{U}_m\right\}.$$

By Urysohn's lemma, for each $(n, m) \in A$ there is some continuous function $f_{n,m} : X \to \mathbb{R}$ such that $f_{n,m}$ is 1 on $\mathcal{U}_m^c$ and 0 on $\overline{\mathcal{U}}_n$ (these are disjoint closed sets). For $x, y \in X$, define

$$\rho(x, y) := \sum_{(n,m) \in A} \frac{1}{2^{n+m}} \left|f_{n,m}(x) - f_{n,m}(y)\right|.$$

The absolute valued term is ≤ 2, so this function will always be finite. Moreover, one can verify that it is indeed a metric on X. It remains to show that it induces the same topology; it suffices to compare bases of the two.

Let $x \in \mathcal{U}_m$. We wish to show there exists $B_\rho(x, \varepsilon) \subseteq \mathcal{U}_m$. $\{x\}$ is closed in X being normal, so there exists some n such that

$$\{x\} \subseteq \mathcal{U}_n \subseteq \overline{\mathcal{U}}_n \subseteq \mathcal{U}_m,$$

so $(n, m) \in A$ and so $f_{n,m}(x) = 0$. Let $\varepsilon = \frac{1}{2^{n+m}}$. Then, if $\rho(x, y) < \varepsilon$, it must be

$$\begin{aligned}
\frac{1}{2^{n+m}} &> \sum_{(n',m') \in A} \frac{1}{2^{n'+m'}} \left|f_{n',m'}(x) - f_{n',m'}(y)\right| \\
&\geq \frac{1}{2^{n+m}} \left|\underbrace{f_{n,m}(x)}_{=0} - f_{n,m}(y)\right| \\
&= \frac{1}{2^{n+m}} \left|f_{n,m}(y)\right|,
\end{aligned}$$

so $\left|f_{n,m}(y)\right| < 1$ and thus $y \notin \mathcal{U}_m^c$ so $y \in \mathcal{U}_m$. It follow that $B_\rho(x, \varepsilon) \subseteq \mathcal{U}_m$, and so every open set in X is open with respect to the metric topology.

Conversely, if $B_\rho(x, \varepsilon)$ some open ball in the metric topology, then notice that $y \mapsto \rho(x, y)$ for fixed y a continuous function, and thus $(\rho(x, \cdot))^{-1}(-\varepsilon, \varepsilon)$ an open set in $\mathcal{T}$ containing x. But this set also just equal to $B_\rho(x, \varepsilon)$, hence $B_\rho(x, \varepsilon)$ open in $\mathcal{T}$. We conclude the two topologies are equal, completing the proof. ■

Remark 1.16: Recall metric ⇒ first countable hence not first countable ⇒ not metrizable.

§1.10 Stone-Weierstrass Theorem

We need to use the following theorem, which we'll prove later.

↪**Theorem 1.10** (Weierstrass Approximation Theorem): Let $f : [a, b] \to \mathbb{R}$ continuous. Then, for every $\varepsilon > 0$, there exists a polynomial $p(x)$ such that $\|f - p\|_\infty < \varepsilon$.

↪**Definition 1.34** (Algebra, Separation of Points): We call a subset $\mathcal{A} \subseteq C(X)$ an *algebra* if it is a linear subspace that is closed under multiplication (that is, $f, g \in \mathcal{A} \Rightarrow f \cdot g \in \mathcal{A}$).

We say $\mathcal{A}$ *separates points* in X if for every $x, y \in X$, there exists an $f \in \mathcal{A}$ such that $f(x) \neq f(y)$.

↪**Theorem 1.11** (Stone-Weierstrass): Let X be a compact Hausdorff space. Suppose $\mathcal{A} \subseteq C(X)$ an algebra that separates points and contains constant functions. Then, $\mathcal{A}$ dense in $C(X)$.

We tacitly assume the conditions of the theorem in the following lemmas as as not to restate them.

↪**Lemma 1.5**: For every $F \subseteq X$ closed, and every $x_0 \in F^c$, there exists a neighborhood $\mathcal{U}(x_0)$ such that $F \cap \mathcal{U} = \varnothing$ and $\forall\, \varepsilon > 0$ there is some $h \in \mathcal{A}$ such that $h < \varepsilon$ on $\mathcal{U}$, $h > 1 - \varepsilon$ on F, and $0 \leq h \leq 1$ on X.

In particular, $\mathcal{U}$ is *independent* of choice of ε.

Proof. Our first claim is that for every $y \in F$, there is a $g_y \in \mathcal{A}$ such that $g_y(x_0) = 0$ and $g_y(y) > 0$, and moreover $0 \leq g_y \leq 1$. Since $\mathcal{A}$ separates points, there is an $f \in \mathcal{A}$ such that $f(x_0) \neq f(y)$. Then, let

$$g_y(x) := \left[\frac{f(x) - f(x_0)}{\|f - f(x)\|_\infty}\right]^2.$$

Then, every operation used in this new function keeps $g_y \in \mathcal{A}$. Moreover one readily verifies it satisfies the desired qualities. In particular since g_y continuous, there is a neighborhood $\mathcal{O}_y$ such that $g_y|_{\mathcal{O}_y} > 0$. Hence, we know that $F \subseteq \bigcup_{y \in F} \mathcal{O}_y$, but F closed and so compact, hence there exists a finite subcover i.e. some $n \geq 1$ and finite sequence $\{y_i\}_{i=1}^n$ such that $F \subseteq \bigcup_{i=1}^n \mathcal{O}_{y_i}$. Let for each y_i $g_{y_i} \in \mathcal{A}$ with the properties from above, and consider the "averaged" function

$$g(x) := \frac{1}{n} \sum_{i=1}^n g_{y_i}(x) \in \mathcal{A}.$$

Then, $g(x_0) = 0$, $g > 0$ on F and $0 \leq g \leq 1$ on all of X. Hence, there is some $1 > c > 0$ such that $g \geq c$ on F, and since g continuous at x_0 there exists some $\mathcal{U}(x_0)$ such that $g < \frac{c}{2}$ on $\mathcal{U}$, with $\mathcal{U} \cap F = \varnothing$. So, $0 \leq g|_{\mathcal{U}} < \frac{c}{2}$, and $1 \geq g|_F \geq c$. To complete the proof, we need $\left(0, \frac{c}{2}\right) \leftrightarrow (0, \varepsilon)$ and $(c, 1) \leftrightarrow (1 - \varepsilon, 1)$. By the Weierstrass Approximation Theorem, there exists some polynomial p such that $p|_{[0, \frac{c}{2}]} < \varepsilon$ and $p|_{[c,1]} > 1 - \varepsilon$. Then if we let $h(x) := (p \circ g)(x)$, this is just a polynomial of g hence remains in $\mathcal{A}$, and we find

$$h|_{\mathcal{U}} < \varepsilon, \qquad h|_F > 1 - \varepsilon, \qquad 0 \leq h \leq 1.$$

■

↪**Lemma 1.6**: For every disjoint closed set A, B and $\varepsilon > 0$, there exists $h \in \mathcal{A}$ such that $h|_A < \varepsilon$, $h|_B > 1 - \varepsilon$, and $0 \leq h \leq 1$ on X.

Proof. Let $F = B$ as in the last lemma. Let $x \in A$, then there exists $\mathcal{U}_x \cap B = \varnothing$ and for every $\varepsilon > 0$, $h|_{\mathcal{U}_x} < \varepsilon$ and $h|_B > 1 - \varepsilon$ and $0 \leq h \leq 1$. Then $A \subseteq \bigcup_{x \in A} \mathcal{U}_x$. Since A closed so compact, $A \subseteq \bigcup_{i=1}^N \mathcal{U}_{x_i}$. Let $\varepsilon_0 < \varepsilon$ such that $\left(1 - \frac{\varepsilon_0}{N}\right)^N > 1 - \varepsilon$. For each i, let $h_i \in \mathcal{A}$ such that $h_i|_{\mathcal{U}_{x_i}} < \frac{\varepsilon_0}{N}$, $h_i|_B > 1 - \frac{\varepsilon_0}{N}$ and $0 \leq h_i \leq 1$. Then, put

$$h(x) = h_1(x) \cdot h_2(x) \cdots h_N(x) \in \mathcal{A}.$$

Then, $0 \leq h \leq 1$ and $h|_B > \left(1 - \frac{\varepsilon_0}{N}\right)^N > 1 - \varepsilon$. Then, for every $x \in A$, $x \in \mathcal{U}_{x_i}$ so $h_i(x) < \frac{\varepsilon_0}{N}$ and $h_i(x) \leq i$ so $h(x) < \frac{\varepsilon_0}{N}$ so $h|A < \frac{\varepsilon_0}{N} < \varepsilon$. ■

Proof. (Of Stone-Weierstrass) WLOG, assume $f \in C(X)$, $0 \leq f \leq 1$, by replacing with

$$\tilde{f}(x) = \frac{f(x) + \|f\|_\infty}{\|f + \|f\|_\infty\|_\infty}$$

if necessary, since if there exists a $\tilde{g} \in \mathcal{A}$ such that $\|\tilde{f} - \tilde{g}\|_\infty < \varepsilon$, then using the properties of $\mathcal{A}$ we can find some appropriate $g \in \mathcal{A}$ such that $\|f - g\|_\infty < \varepsilon$.

Fix $n \in \mathbb{N}$, and consider the set $\left\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\right\}$, and let for $1 \leq j \leq n$

$$A_j := \left\{x \in X \mid f(x) \leq \frac{j-1}{n}\right\}, \qquad B_j := \left\{x \in X \mid f(x) \geq \frac{j}{n}\right\},$$

which are both closed and disjoint. By the lemma, there exists $g_j \in \mathcal{A}$ such that

$$g_j|_{A_j} < \frac{1}{n}, \qquad g_j|_{B_j} > 1 - \frac{1}{n},$$

with $0 \leq g_j \leq 1$. Let then

$$g(x) := \frac{1}{n} \sum_{j=1}^n g_j(x) \in \mathcal{A}.$$

We claim then $\|f - g\|_\infty \leq \frac{3}{n}$, which proves the claim by taking n sufficiently large.

Suppose $k \in [1, n]$. If $f(x) \leq \frac{k}{n}$, then

$$g_j(x) = \begin{cases} < \frac{1}{n} \text{ if } j-1 \geq k \\ \leq 1 \text{ else} \end{cases},$$

so

$$g(x) = \frac{1}{n}\sum_{j=1}^{n} g_j(x) = \frac{1}{n}\left[\sum_{j=1}^{k} g_j(x) + \sum_{j=k+1}^{n} g_j(x)\right] \leq \frac{1}{n}\left[k + \frac{n-k}{n}\right] \leq \frac{k}{n} + \frac{n-k}{n^2} \leq \frac{k+1}{n}.$$

Similarly if $f(x) \geq \frac{k-1}{n}$, then

$$g_j(x) = \begin{cases} > 1 - \frac{1}{n} \text{ if } j \leq k-1 \\ \geq 0 \quad\quad \text{else} \end{cases},$$

so

$$g(x) \geq \frac{1}{n}\sum_{j=1}^{k-1}\left(1 - \frac{1}{n}\right) \geq \frac{1}{n}(k-1)\left(1 - \frac{1}{n}\right) = \frac{k-1}{n} - \frac{k-1}{n^2} \geq \frac{k-2}{n}.$$

So, we've show that if $\frac{k-1}{n} \leq f(x) \leq \frac{k}{n}$, then $\frac{k-2}{n} \leq g(x) \leq \frac{k+1}{n}$, and so repeating this argument and applying triangle inequality we conclude $\|f-g\|_\infty \leq \frac{3}{n}$. ■

↪**Theorem 1.12** (Borsuk): X compact, Hausdorff and $C(X)$ separable ⇔ X is metrizable.

§2 Functional Analysis

Here, we will primarily work with a normed vector space (nvs). Moreover, we usually work in:

↪**Definition 2.1** (Banach Space): A normed vector space $(X, \|\cdot\|)$ is a *Banach space* if it is complete as a metric space under the norm-induced metric.

§2.1 Introduction to Linear Operators

↪**Definition 2.2** (Linear Operator, Operator Norm): Let X, Y be vector spaces. Then, a map $T : X \to Y$ is called *linear* if $\forall\, x, y \in X, \alpha, \beta \in \mathbb{R}, T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$.

If X, Y normed vector spaces, we say T is a bounded linear operator if T linear and the *operator norm*

$$\|T\| = \|T\|_{\mathcal{L}(X,Y)} = \sup_{\substack{x \in X, \\ \|x\|_X \leq 1}} \|Tx\|_Y < \infty$$

is finite. Then, we put

$$\mathcal{L}(X,Y) := \{\text{bounded linear operators } X \to Y\}.$$

We'll also write $\mathcal{L}(X) := \mathcal{L}(X,X)$.

↪**Theorem 2.1** (Bounded iff Continuous): If X, Y are nvs, $T \in \mathcal{L}(X, Y)$ iff and only if T is continuous, i.e. if $x_n \to x$ in X, then $Tx_n \to Tx$ in Y.

Proof. If $T \in \mathcal{L}(X, Y)$,

$$\begin{aligned} \|Tx_n - Tx\|_Y &= \|T(x_n - x)\|_Y \\ &= \|x_n - x\|_X \cdot \|\frac{T(x_n - x)}{\|x_n - x\|_X}\|_Y \\ &\leq \underbrace{\|T\|}_{<\infty} \|x_n - x\|_X \to 0, \end{aligned}$$

hence T continuous. Conversely, if T continuous, then by linearity $T0 = 0$, so by continuity, there is some $\delta > 0$ such that $\|Tx\|_Y < 1$ if $\|x\|_X < \delta$. For $x \in X$ nonzero, let $\lambda = \frac{\delta}{\|x\|_X}$. Then, $\|\lambda x\|_X \leq \delta$ so $\|T(\lambda x)\|_Y < 1$, i.e. $\frac{\| T(x)\|_Y \delta}{\|x\|_X} < 1$. Hence,

$$\|T\| = \sup_{x \in X : x \neq 0} \frac{\| T(x)\|_Y}{\|x\|_X} \leq \frac{1}{\delta},$$

so $T \in \mathcal{L}(X, Y)$. ■

↪**Proposition 2.1** (Properties of $\mathcal{L}(X, Y)$): If X, Y nvs, $\mathcal{L}(X, Y)$ a nvs, and if X, Y Banach, then so is $\mathcal{L}(X, Y)$.

Proof. (a) For $T, S \in \mathcal{L}(X, Y)$, $\alpha, \beta \in \mathbb{R}$, and $x \in X$, then

$$\begin{aligned} \|(\alpha T + \beta S)(x)\|_Y &\leq |\alpha| \ \|Tx\|_Y + |\beta| \ \|Sx\|_Y \\ &\leq |\alpha| \ \|T\| \ \|x\|_X + |\beta| \ \|T\| \ \|x\|_X. \end{aligned}$$

Dividing both sides by $\|x\|$, we find $\|\alpha T + \beta S\| < \infty$. The same argument gives the triangle inequality on $\| \cdot \|$. Finally, $T = 0$ iff $\|Tx\|_Y = 0$ for every $x \in X$ iff $\|T\| = 0$.

(b) Let $\{T_n\} \subseteq \mathcal{L}(X, Y)$ be a Cauchy sequence. We have that

$$\|T_n x - T_m x\|_Y \leq \|T_n - T_m\| \ \|x\|_X,$$

so in particular the sequence $\{T_n(x)\}$ a Cauchy sequence in Y for any $x \in X$. Y complete so this sequence converges, say $T_n(x) \to y^*$ in Y. Let $T(x) := y^*$ for each x. We claim that $T \in \mathcal{L}(X, Y)$ and that $T_n \to T$ in the operator norm. We check:

$$\begin{aligned} \alpha T(x_1) + \beta T(x_2) &= \lim_{n \to \infty} \alpha T_n(x_1) + \lim_{n \to \infty} \beta T_n(x_2) \\ &= \lim_{n \to \infty} [T_n(\alpha x_1) + T_n(\beta x_2)] \\ &= \lim_{n \to \infty} T_n(\alpha x_1 + \beta x_2) \\ &= T(\alpha x_1 + \beta x_2), \end{aligned}$$

so T linear.

Let now $\varepsilon > 0$ and N such that for every $n \geq N$ and $k \geq 1$ such that $\|T_n - T_{n+k}\| < \frac{\varepsilon}{2}$. Then,

$$\begin{aligned}\|T_n(x) - T_{n+k}(x)\|_Y &= \left\|(T_n - T_{n+k})(x)\right\|_Y \\ &\leq \left\|T_n - T_{n+k}\right\| \|x\|_X \\ &< \frac{\varepsilon}{2} \|x\|_X.\end{aligned}$$

Letting $k \to \infty$, we find that

$$\|T_n(x) - T(x)\|_Y < \frac{\varepsilon}{2}\ \|x\|_X,$$

so normalizing both sides by $\|x\|_X$, we find $\|T_n - T\| < \frac{\varepsilon}{2}$, and we have convergence. ■

↪**Definition 2.3** (Isomorphism): We say $T \in \mathcal{L}(X, Y)$ an *isomorphism* if T is bijective and $T^{-1} \in \mathcal{L}(Y, X)$. In this case we write $X \simeq Y$, and say X, Y isomorphic.

§2.2 Finite versus Infinite Dimensional

If X a nvs, then we can look for a basis β such that $\text{span}(\beta) = X$. If $\beta = \{e_1, ..., e_n\}$ has no proper subset spanning X, then we say $\dim(X) = n$.

As we saw on homework, any two norms on a finite dimensional space are equivalent.

↪**Corollary 2.1**: (a) Any two nvs of the same finite dimension are isomorphic.

(b) Any finite dimensional space is complete, and so any finite dimensional subspace is closed.

(c) $\overline{B}(0, 1)$ is compact in a finite dimensional space.

PROOF. (a) Let $(X, \|\cdot\|)$ have finite dimension n. Then, we claim $(X, \|\cdot\|) \simeq (\mathbb{R}^n, |\cdot|)$. Let $\{e_1, ..., e_n\}$ be a basis for X. Let $T : \mathbb{R}^n \to X$ given by

$$T(x) = \sum_{i=1}^{n} x_i e_i,$$

where $x = (x_1, ..., x_n) \in \mathbb{R}^n$, which is clearly linear. Moreover,

$$Tx = 0 \Leftrightarrow \sum_{i=1}^{n} x_i e_i = 0 \Leftrightarrow x = 0,$$

so T injective, and so being linear between two spaces of the same dimension gives T surjective. It remains to check boundedness.

First, we claim $x \mapsto \|T(x)\|$ is a norm on $\mathbb{R}^n$. $\|T(x)\| = 0 \Leftrightarrow x = 0$ by the injectivity of T, and the properties $\|T(\lambda x)\| = |\lambda|\ \|Tx\|$ and $\|T(x + y)\| \leq \|Tx\| + \|Ty\|$ follow from linearity of T and the fact that $\|\cdot\|$ already a norm. Hence, $\|T(\cdot)\|$ a norm on $\mathbb{R}^n$ and so equivalent to $|\cdot|$, i.e. there exists constants $C_1, C_2 > 0$ such that

$$C_1|x| \leq \|T(x)\| \leq C_2|x|,$$

for every $x \in X$. It follows that $\|T\|$ (operator norm now) is bounded.

Letting $T(x) = y$, we find similarly

$$C_{1'}\|y\| \leq |T^{-1}(y)| \leq C_{2'}\ \|y\|,$$

so $\|T^{-1}\|$ also bounded. Hence, we've shown any n-dimensional space is isomorphic to $\mathbb{R}^n$, so by transitivity of isomorphism any two n-dimensional spaces are isomorphic.

(b) The property of completeness is preserved under isomorphism, so this follows from the previous statement since $\mathbb{R}^n$ complete.

(c) Consider $\overline{B}(0,1) \subseteq X$. Let T be an isomorphism $X \to \mathbb{R}^n$. Then, for $x \in \overline{B}(0,1)$, $\|Tx\| \leq \|T\| < \infty$, so $T\big(\overline{B}(0,1)\big)$ is a bounded subset of $\mathbb{R}^n$, and since T and its inverse continuous, $T\big(\overline{B}(0,1)\big)$ closed in $\mathbb{R}^n$. Hence, $T\big(\overline{B}(0,1)\big)$ closed and bounded hence compact in $\mathbb{R}^n$, so since T^{-1} continuous $T^{-1}\Big(T\big(\overline{B}(0,1)\big)\Big) = \overline{B}(0,1)$ also compact, in X. ■

↪**Theorem 2.2** (Riesz's): If X is an nvs, then $\overline{B}(0,1)$ is compact if and only if X if finite dimensional.

↪**Lemma 2.1** (Riesz's): Let $Y \subsetneq X$ be a closed nvs (and X a nvs). Then for every $\varepsilon > 0$, there exists $x_0 \in X$ with $\|x_0\| = 1$ and such that

$$\|x_0 - y\|_X > \varepsilon \,\forall\, y \in Y.$$

Proof. Fix $\varepsilon > 0$. Since $Y \subsetneq X$, let $x \in Y^c$. Y closed so Y^c open and hence there exists some $r > 0$ such that $B(x,r) \cap Y = \varnothing$. In other words,

$$\inf\{\|x - y'\| \mid y' \in Y\} > r > 0.$$

Let then $y_1 \in Y$ be such that

$$r < \|x - y_1\| < \varepsilon^{-1} r,$$

and take

$$x_0 := \frac{x - y_1}{\|x - y_1\|_X}.$$

Then, x_0 a unit vector, and for every $y \in Y$,

$$\begin{aligned} x_0 - y &= \frac{x - y_1}{\|x - y_1\|} - y \\ &= \frac{1}{\|x - y_1\|}[x - y_1 - y\ \|x - y_1\|] \\ &= \frac{1}{\|x - y_1\|}[x - y'], \end{aligned}$$

where $y' = y_1 + y\,\|x - y_1\| \in Y$, since it is closed under vector addition. Hence

$$\|x_0 - y\| = \frac{1}{\|x - y_1\|}\,\|x - y'\| > \frac{\varepsilon}{r}\,\|x - y'\| > \varepsilon,$$

for every $y \in Y$. ■

PROOF. (Of Thm. 2.2) ($\Leftarrow$) By the previous corollary.

($\Rightarrow$) Suppose X infinite dimensonal. We will show $B := \overline{B}(0,1)$ not compact.

Claim: there exists $\{x_i\}_{i=1}^{\infty} \subseteq B$ such that $\|x_i - x_j\| > \frac{1}{2}$ if $i \neq j$.

We proceed by induction. Let $x_1 \in B$. Suppose $\{x_1, ..., x_n\} \subseteq B$ are such that $\|x_i - x_j\| > \frac{1}{2}$. Let $X_n = \mathrm{span}\{x_1, ..., x_n\}$, so X_n finite dimensional hence $X_n \subsetneq X$. By the previous lemma (taking $\varepsilon = \frac{1}{2}$) there is then some $x_{n+1} \in B$ such that $\|x_1 - x_{n+1}\| > \frac{1}{2}$ for every $i = 1, ..., n$. We can thus inductively build such a sequence $\{x_i\}_{i=1}^{\infty}$. Then, every subsequence of this sequence cannot be Cauchy so B is not sequentially compact and thus B is not compact. ■

§2.3 Open Mapping and Closed Graph Theorems

↪**Definition 2.4** (T open)**:** If X, Y toplogical spaces and $T : X \to Y$ a linear operator, T is said to be *open* if for every $\mathcal{U} \subseteq X$ open, $T(\mathcal{U})$ open in Y.

In particular if X, Y are metric spaces (or nvs), then T is open iff the image of every open ball in X containes an open ball in Y, i.e. $\forall\, x \in X, r > 0$ there exists $r' > 0$ such that $T(B_X(x, r)) \supseteq B_Y(Tx, r')$. Moreover, by translating/scaling appropriately, it suffices to prove for $x = 0, r = 1$.

↪**Theorem 2.3** (Open Mapping Theorem)**:** Let X, Y be Banach spaces and $T : X \to Y$ a bounded linear operator. If T is surjective, then T is open.

PROOF. Its enough to show that there is some $r > 0$ such that $T(B_X(0,1)) \supseteq B_Y(0, r)$.

Claim: $\exists\, c > 0$ such that $\overline{T(B_X(0,1))} \supseteq B_Y(0, 2c)$.

Put $E_n = n \cdot \overline{T(B_X(0,1))}$ for $n \in \mathbb{N}$. Since T surjective, $\bigcup_{n=1}^{\infty} E_n = Y$. Each E_n closed, so by the Baire Category Theorem there exists some index n_0 such that E_{n_0} has nonempty interior, i.e.

$$\mathrm{int}\left(\overline{T(B_X(0,1))}\right) \neq \varnothing,$$

where we drop the index by homogeneity. Pick then $c > 0$ and $y_0 \in Y$ such that $B_Y(y_0, 4c) \subseteq \overline{T(B_X(0,1))}$. We claim then that $B_Y(-y_0, 4c) \subseteq \overline{T(B_X(0,1))}$ as well. Indeed, if $B_Y(y_0, 4c) \subseteq \overline{T(B_X(0,1))}$, then $\forall\, \tilde{y} \in Y$ with $\|y_0 - \tilde{y}\|_Y < 4c$, Then, $\|-y_0 + \tilde{y}\|_Y < 4c$ so $-\tilde{y} \in B_Y(-y_0, 4c)$. But $\tilde{y} = \lim_{n\to\infty} T(x_n)$ and so $-\tilde{y} = \lim_{n\to\infty} T(-x_n)$. Since $\{-x_n\} \subseteq B_X(0,1)$, this implies $-\tilde{y} \in \overline{T(B_X(0,1))}$ hence the "subclaim" holds.

Now, for any $\tilde{y} \in B_Y(0, 4c)$, $\|\tilde{y}\| \leq 4c$ so

$$\tilde{y} = y_0 \underbrace{-y_0 + \tilde{y}}_{\in B_Y(-y_0, 4c)} = \overbrace{y_0 + \tilde{y}}^{\in B(y_0, 4c)} - y_0.$$

Therefore,

$$\begin{aligned} B_Y(0, 4c) &= B_Y(y_0 - y_0, 4c) \\ &\subseteq B_Y(y_0, 4c) + B_Y(-y_0, 4c) \\ &\overline{T(B_X(0,1))} + \overline{T(B_X(0,1))} = 2\overline{T(B_X(0,1))}, \end{aligned}$$

(where summation of two sets is the vector addition of all the elements in the sets), hence $B_Y(0, 2c) \subseteq \overline{T(B_X(0,1))}$.

We claim next that $T(B_X(0,1)) \supseteq B_Y(0,c)$. Choose $y \in Y$ with $\|y\|_Y < c$. By the first claim, $B_Y(0,c) \subseteq \overline{T(B_X(0, \frac{1}{2}))}$, so for every $\varepsilon > 0$ there is some $z \in X$ with $\|z\|_X < \frac{1}{2}$ and $\|y - Tz\|_Y < \varepsilon$. Let $\varepsilon = \frac{c}{2}$ and $z_1 \in X$ such that $\|z_1\|_X < \frac{1}{2}$ and $\|y - Tz_1\|_Y < \frac{c}{2}$. But the first claim can also be written as $B_Y(0, \frac{c}{2}) \subseteq \overline{T(B_X(0, \frac{1}{4}))}$ so if $\varepsilon = \frac{c}{4}$, let $z_2 \in X$ such that $\|z_2\|_X < \frac{1}{4}$ and $\|(y - Tz_1) - Tz_2\|_Y < \frac{c}{4}$. Continuing in this manner we find that

$$B_Y\left(0, \frac{c}{2^k}\right) \subseteq \overline{T\left(B_X\left(0, \frac{1}{2^{k+1}}\right)\right)},$$

so exists $z_k \in X$ such that $\|z_k\|_X < \frac{1}{2^k}$ and $\|y - T(z_1 + \cdots + z_k)\|_Y < \frac{c}{2^k}$. Let $x_n = z_1 + \cdots + z_n \in X$. Then $\{x_n\}$ is Cauchy in X, since

$$\|x_n - x_m\|_X \leq \sum_{k=m}^{n} \|z_k\|_X < \sum_{k=m}^{n} \frac{1}{2^k} \to 0.$$

Since X a Banach space, $x_n \to \overline{x}$ and in particular $\|\overline{x}\| \leq \sum_{k=1}^{\infty} \|z_k\|_X < \sum_{k=1}^{\infty} \frac{1}{2^k} = 1$, so $\overline{x} \in B_X(0,1)$. Since T bounded it is continuous, so $Tx_n \to T\overline{x}$, so $y = T\overline{x}$ and thus $B_Y(0,c) \subseteq T(B(0,1))$. ■

↪**Corollary 2.2**: Let X, Y Banach and $T : X \to Y$ be bounded, linear and bijective. Then, T^{-1} continuous.

Proof. Let $\mathcal{U} \subseteq X$ open. Then, $\left(T^{-1}\right)^{-1}(\mathcal{U}) = T(\mathcal{U})$ is open since T surjective, so T^{-1} continuous. ■

↪**Corollary 2.3**: Let $(X, \|\cdot\|_1), (X, \|\cdot\|_2)$ be Banach spaces. Suppose there exists $c > 0$ such that $\|x\|_2 \leq C\|x\|_1$ for every $x \in X$. Then, $\|\cdot\|_1, \|\cdot\|_2$ are equivalent.

Proof. Let T be the identity linear operator and use the previous corollary. ■

↪**Definition 2.5** (T closed): If X, Y are nvs and T is linear, the *graph* of T is the set

$$G(T) = \{(x, Tx) \mid x \in X\} \subseteq X \times Y.$$

We then say T is *closed* if $G(T)$ closed in $X \times Y$.

Remark 2.1: Since X, Y are nvs, they are metric spaces so first countable, hence closed $\leftrightarrow$ contains all limit points.

In the product topology, a countable base for $X \times Y$ at (x, y) is given by

$$\left\{B_X\left(x, \frac{1}{n}\right) \times B\left(y, \frac{1}{m}\right)\right\}_{n,m\in\mathbb{N}}.$$

Then, $G(T)$ closed iff $G(T)$ contains all limit points. How can we put a norm on $X \times Y$ that generates this product topology? Let

$$\|(x, y)\|_1 := \|x\|_X + \|y\|_Y.$$

If $(x_n, y_n) \to (x, y)$ in the product topology, then since Π_1, Π_2 continuous maps, $(x_n, y_n) \to (x, y)$ in the $\|\cdot\|_1$ topology. On the other hand if $(x_n, y_n) \to (x, y)$ in the $\|\cdot\|_1$ norm, then

$$\|x_n - x\|_X \leq \|(x_n, y_n) - (x, y)\|_1,$$

hence since the RHS $\to 0$ so does the LHS and so $x_n \to x$ in $\|\cdot\|_X$; similar gives $y_n \to y$ in $\|\cdot\|_Y$. From here it follows that $(x_n, y_n) \to (x, y)$ in the product topology.

So, to prove $G(T)$ closed, we just need to prove that if $x_n \to x$ in X and $Tx_n \to y$, then $y = Tx_n$.

↪**Theorem 2.4** (Closed Graph Theorem): Let X, Y be Banach spaces and $T : X \to Y$ linear. Then, T is continuous iff T is closed.

Proof. ($\Rightarrow$) Immediate from the above remark.

($\Leftarrow$) Consider the function

$$x \mapsto \|x\|_* := \|x\|_X + \|Tx\|_Y.$$

So by the above, T closed implies $(X, \|\cdot\|_*)$ is complete, i.e. if $x_n \to x$ in $\|\cdot\|_*$ in X iff $x_n \to x$ in $\|\cdot\|_X$ and $Tx_n \to Tx$ in $\|\cdot\|_Y$. However, $\|\cdot\|_X \leq \|\cdot\|_*$, hence since $\left(X, \|\cdot\|_X\right)$ and $\left(X, \|\cdot\|_*\right)$ are Banach spaces, by the corollary, there is some $C > 0$ such that $\|\cdot\|_* \leq C\|\cdot\|_X$. So,

$$\|x\|_X + \|Tx\|_Y \leq C\|x\|_X,$$

so

$$\|Tx\|_Y \leq \|x\|_X + \|Tx\|_Y \leq C\|x\|_X,$$

so T bounded and thus continuous. ■

Remark 2.2: The Closed Graph Theorem simplifies proving continuity of T. It tells us we can assume if $x_n \to x$, $\{Tx_n\}$ Cauchy so $\exists\, y$ such that $Tx_n \to y$ since Y is Banach. So, it suffices to check that $y = Tx$ to check continuity; we don't need to check convergence of Tx_n.

§2.4 Uniform Boundedness Principle

Recall the following consequence of the Baire Category Theorem:

↪**Theorem 2.5**: Let $\mathcal{F} \subseteq C(X)$ where (X, ρ) a complete metric space. Suppose $\mathcal{F}$ pointwise bounded. Then, there exists a nonempty open set $\mathcal{O} \subseteq X$ such that there is some $M > 0$ such that $|f(x)| \leq M$ for every $x \in \mathcal{O}$, $f \in \mathcal{F}$.

This leads to the following result:

↪**Theorem 2.6** (Uniform Boundedness Principle): Let X a Banach space and Y a nvs. Consider $\mathcal{F} \subseteq \mathcal{L}(X, Y)$. Suppose $\mathcal{F}$ is pointwise bounded, i.e. for every $x \in X$, there is some $M_x > 0$ such that

$$\|Tx\|_Y \leq M_x, \forall\, T \in \mathcal{F}.$$

Then, $\mathcal{F}$ is uniformly bounded, i.e. $\exists\, M > 0$ such that

$$\|T\|_Y \leq M, \forall\, T \in \mathcal{F}.$$

Proof. For every $T \in \mathcal{F}$, let $f_T : X \to \mathbb{R}$ be given by

$$f_T(x) = \|Tx\|_Y.$$

Since $T \in \mathcal{L}(X, Y)$, T is continuous, so $x_n \underset{X}{\to} x \Rightarrow Tx_n \underset{Y}{\to} Tx$, hence $\|Tx_n\|_Y \to \|Tx\|_Y$ so f_T continuous for each T i.e. $f_T \in C(X)$, so $\{f_T\} \subseteq C(X)$ pointwise bounded. So by the previous theorem, there is some ball $B(x_0, r) \subseteq X$ and some $K > 0$ such that $\|Tx\| \leq K$ for every $x \in B(x_0, r)$ and $T \in \mathcal{F}$. Thus, for every $x \in B(0, r)$,

$$\begin{aligned} \|Tx\| &= \|T(x - x_0 + x_0)\| \\ &\leq \Bigg\| T \underbrace{(x - x_0)}_{\in B(x_0, r)} \Bigg\| + \|Tx_0\| \\ &\leq K + M_{x_0}, \qquad \forall\, x \in B(0, r), T \in \mathcal{F}. \end{aligned}$$

Thus, for every $x \in B(0, 1)$,

$$\|Tx\| = \frac{1}{r} \Bigg\| T \underbrace{(rx)}_{\in B(0, r)} \Bigg\| \leq \frac{1}{r}\left(K + M_{x_0}\right) =: M,$$

so its clear $\|T\| \leq M$ for every $T \in \mathcal{F}$. ■

↪**Theorem 2.7** (Banach-Saks-Steinhaus): Let X a Banach space and Y a nvs. Let $\{T_n\} \subseteq \mathcal{L}(X, Y)$. Suppose for every $x \in X$, $\lim_{n\to\infty} T_n(x)$ exists in Y. Then,

a. $\{T_n\}$ are uniformly bounded in $\mathcal{L}(X, Y)$;

b. For $T : X \to Y$ defined by $T(x) := \lim_{n\to\infty} T_n(x)$, we have $T \in \mathcal{L}(X, Y)$;

c. $\|T\| \leq \liminf_{n\to\infty} \|T_n\|$ (*lower semicontinuity result*).

Proof. (a) For every $x \in X$, $T_n(x) \to T(x)$ so $\|Tx\| < \infty$ hence $\sup_n \|T_n x\| < \infty$. By uniform boundedness, then, we find $\sup_n \|T_n\| =: C < \infty$.

(b) T is linear (by linearity of T_n). By (a),

$$\|T_n x\| \leq C\|x\|,$$

for every n, x, so

$$\|Tx\| \leq C\|x\| \,\forall\, x \in X,$$

so T bounded.

(c) We know

$$\|T_n x\| \leq \|T_n\|\|x\| \,\forall\, x \in X,$$

so

$$\frac{\|T_n x\|}{\|x\|} \leq \|T_n\|,$$

so

$$\liminf_n \frac{\|T_n x\|}{\|x\|} = \frac{\|Tx\|}{\|x\|} \leq \liminf_n \|T_n\|,$$

so by "suping" both sides,

$$\|T\| \leq \liminf_n \|T_n\|.$$

■

Remark 2.3:

- We do not necessarily have $T_n \to T$ in $\mathcal{L}(X, Y)$ i.e. with respect to the operator norm.
- If Y is a Banach space, then $\lim_{n\to\infty} T_n(x)$ exists in $Y \Leftrightarrow \{T_n x\}$ Cauchy in Y for every $x \in X$.

§2.5 Introduction to Hilbert Spaces

↪**Definition 2.6** (Inner Product): An *inner product* on a vector space X is a map $(\cdot,\cdot) : X \times X \to \mathbb{R}$ such that for every $\lambda, \mu \in \mathbb{R}$ and $x, y, z \in X$,

- $(\lambda x + \mu y, z) = \lambda(x, z) + \mu(y, z)$;
- $(x, y) = (y, x)$;
- $(x, x) \geq 0$ and $(x, x) = 0 \Leftrightarrow x = 0$.

Remark 2.4: The first and second conditions combined imply that $(\cdot,\cdot)$ actually *bilinear*, namely, linear in both coordinates.

Remark 2.5: An inner product induces a norm on a vector space by

$$\|x\| := (x, x)^{\frac{1}{2}}.$$

↪**Proposition 2.2** (Cauchy-Schwarz Inequality): Any inner product satisfies Cauchy-Schwarz, namely,

$$|(x, y)| \leq \|x\|\|y\|,$$

for every $x, y \in X$.

Proof. Suppose first $y = 0$. Then, the right hand side is clearly 0, and by linearity $(x, y) = 0$, hence we have $0 \leq 0$ and are done. Suppose then $y \neq 0$. Then, let $z = x - \frac{(x,y)}{(y,y)}y$ where $y \neq 0$. Then,

$$\begin{aligned}
0 \leq \|z\|^2 &= \left(x - \frac{(x, y)}{(y, y)}y, x - \frac{(x, y)}{(y, y)}y\right) \\
&= (x, x) - \frac{(x, y)}{(y, y)}(x, y) - \frac{(x, y)}{(y, y)}(y, x) + \frac{(x, y)^2}{(y, y)^2}(y, y) \\
&= (x, x) - \frac{2((x, y))^2}{(y, y)} + \frac{(x, y)^2}{(y, y)} \\
&= \|x\| - \frac{(x, y)^2}{(y, y)} \\
&\Rightarrow \frac{(x, y)^2}{(y, y)} \leq \|x\| \Rightarrow (x, y)^2 \leq \|x\|^2\|y\|^2 \\
&\Rightarrow |(x, y)| \leq \|x\|\|y\|.
\end{aligned}$$

■

↪**Corollary 2.4**: The function $\|x\| := (x, x)^{\frac{1}{2}}$ is actually a norm on X.

PROOF. By definition, $\|x\| \geq 0$ and equal to zero only when $x = 0$. Also,

$$\|\lambda x\| = (\lambda x, \lambda x)^{\frac{1}{2}} = |\lambda|(x,x)^{\frac{1}{2}} = |\lambda|\|x\|.$$

Finally,

$$\begin{aligned}
\|x+y\|^2 &= (x+y, x+y) \\
&= (x,x) + 2(x,y) + (y,y) \\
&= \|x\|^2 + \|y\|^2 + 2(x,y) \\
\text{by Cauchy-Schwarz} \qquad &\leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| \\
&= (\|x\| + \|y\|)^2,
\end{aligned}$$

hence by taking square roots we see $\|x+y\| \leq \|x\| + \|y\|$ as desired. ■

↪**Proposition 2.3** (Parallelogram Law): Any inner product space satisfies the following:

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

↪**Corollary 2.5**: $(\cdot,\cdot)$ is continuous, i.e. if $x_n \to x$ and $y_n \to y$, then $(x_n, y_n) \to (x,y)$.

PROOF.

$$\begin{aligned}
|(x_n,y_n) - (x,y)| &= |(x_n,y_n) - (x,y_n) + (x,y_n) - (x,y)| \\
&= |(x_n - x, y_n) + (x, y_n - y)| \\
&\leq |(x_n - x, y_n)| + |(x, y_n - y)| \\
\text{(Cauchy-Schwarz)} \qquad &\leq \underbrace{\|x_n - x\|}_{\to 0} \underbrace{\|y_n\|}_{\leq M} + \|x\| \underbrace{\|y_n - y\|}_{\to 0} \to 0.
\end{aligned}$$

■

↪**Definition 2.7** (Hilbert Space): A *Hilbert Space* H is a complete inner product space, namely, it is complete with respect to the norm induced by the inner product.

⊛ **Example 2.1**:

1. ℓ^2, the space of square-summable real-valued sequences, equipped with inner product $(x,y) = \sum_{i=1}^{\infty} x_i y_i$.
2. L^2, with inner product $(f,g) = \int f(x)g(x)\,\mathrm{d}x$.

↪**Definition 2.8** (Orthogonality): We say x, y *orthogonal* and write $x \perp y$ if $(x, y) = 0$. If $M \subseteq H$, then the *orthogonal complement* of M, denoted $M^\perp$, is the set

$$M^\perp = \{y \in H \mid (x, y) = 0, \forall\, x \in M\}.$$

Remark 2.6: $M^\perp$ is always a closed subspace of H. If $y_1, y_2 \in M^\perp$, then for every $x \in M$,

$$(x, \alpha y_1 + \beta y_2) = \alpha(x, y_1) + \beta(x, y_2) = 0,$$

so $M^\perp$ a subspace.

If $y_n \to y$ in the norm on H and $\{y_n\} \subseteq M^\perp$, then using the continuity of $(\cdot, \cdot)$, we know that for every $x \in M$, $(x, y_n) \to (x, y)$. But the $(x, y_n) = 0$ for every n and thus $(x, y) = 0$ so $y \in M^\perp$, hence $M^\perp$ closed.

↪**Proposition 2.4**: If $M \subsetneq H$ is a closed subspace, then every $x \in H$ has a unique decomposition

$$x = u + v, \qquad u \in M, v \in M^\perp.$$

Hence, we may write $H = M \oplus M^\perp$. Moreover,

$$\|x - u\| = \inf_{y \in M} \|x - y\|, \qquad \|x - v\| = \inf_{y \in M^\perp} \|x - y\|.$$

Proof. Let $x \in H$. If $x \in M$, we're done with $u = x, v = 0$. Else, if $x \notin M$, then we claim that there is some $u \in M$ such that $\|x - u\| = \inf_{y \in M} \|x - y\| =: \delta > 0$. By definition of the infimum, there exists a sequence $\{u_n\} \subseteq M$ such that

$$\|x - u_n\|^2 \leq \delta^2 + \frac{1}{n}.$$

Let $\overline{x} := u_m - x$, $\overline{y} = u_n - x$. By the Parallelogram Law,

$$\|\overline{x} - \overline{y}\|^2 + \|\overline{x} + \overline{y}\|^2 = 2\|\overline{x}\|^2 + 2\|\overline{y}\|^2$$

hence

$$\|u_m - u_n\|^2 + \|u_m + u_n - 2x\|^2 = 2\|u_m - x\|^2 + 2\|u_n - x\|^2.$$

Now, the second term can be written

$$\|u_m + u_n - 2x\|^2 = 4\left\|\frac{u_m + u_n}{2} - x\right\|^2,$$

hence we find

$$\|u_m - u_n\|^2 = 2\|u_m - x\|^2 + 2\|u_n - x\|^2 - 4\left\|\frac{u_m + u_n}{2} - x\right\|^2.$$

Recall that M a subspace, hence $\frac{1}{2}(u_m + u_n) \in M$ so $\left\|x - \frac{1}{2}(u_m + u_n)\right\| \geq \delta$ as defined before. Thus, we find that by our choice of $\{u_n\}$,

$$\|u_m - u_n\|^2 \leq 2\left(\delta^2 + \frac{1}{m}\right) + 2\left(\delta^2 + \frac{1}{n}\right) - 4\delta^2 = \frac{2}{m} + \frac{2}{n},$$

and thus, by making m, n sufficiently large we can make $\|u_m - u_n\|$ arbitrarily small. Hence, $\{u_n\} \subseteq M$ are Cauchy. H is complete, hence the $\{u_n\}$'s converge, and thus since M closed, $u_n \to u \in M$. Then, we find

$$\begin{aligned} \|x - u\| &\leq \|x - u_n\| + \|u_n - u\| \\ &\leq \underbrace{\left(\delta^2 + \frac{1}{n}\right)^{\frac{1}{2}}}_{\to \delta} + \underbrace{\|u_n - u\|}_{\to 0} \to \delta. \end{aligned}$$

But also, $u \in M$ and thus $\|x - y\| \geq \delta$, and we conclude $\|x - u\| = \delta = \inf_{y \in M}\|x - y\|$.

Next, we claim that if we define $v = x - y$, then $v \in M^{\perp}$. Consider $y \in M$, $t \in \mathbb{R}$, then

$$\Bigg\|x - \underbrace{(u - ty)}_{\in M}\Bigg\|^2 = \|v + ty\|^2 = \|v\|^2 + 2t(v, y) + t^2\|y\|^2.$$

Then, notice that the map

$$t \mapsto \|v + ty\|^2$$

is minimized when $t = 0$, since $\|x - z\|$ for $z \in M$ is minimized when $z = u$, as we showed in the previous part, so equivalently $\|x - (u - ty)\|^2$ minimized when $t = 0$. Thus,

$$\begin{aligned} 0 = \frac{\partial}{\partial t}\|v + ty\|^2|_{t=0} &= \frac{\partial}{\partial t}\left[\|v\|^2 + 2t(v, y) + t^2\|y\|^2\right]_{t=0} \\ &= \left(2(v, y) + 2t\|y\|^2\right)_{t=0} = (v, y) \\ &\Rightarrow (v, y) = 0 \,\forall\, y \in M \Rightarrow v \in M^{\perp}. \end{aligned}$$

So, $x = u + v$ and $u \in M, v \in M^{\perp}$. For uniqueness, suppose $x = u_1 + v_1 = u_2 + v_2$. Then, $u_1 - u_2 = v_2 - v_1$, but then

$$\|v_2 - v_1\|^2 = (v_2 - v_1, v_2 - v_1) = (v_2 - v_1, u_2 - u_1) = 0,$$

so $v_2 = v_1$ so it follows $u_2 = u_1$ and uniqueness holds. ■

↪**Definition 2.9** (Dual of H): The *dual* of H, denoted H^*, is the set

$$H^* := \{f : H \to \mathbb{R} \mid f \text{ continuous and linear}\}.$$

On this space, we may equip the operator norm

$$\|f\|_{H^*} = \|f\| = \sup_{x \in H} \frac{|f(x)|}{\|x\|_H} = \sup_{\|x\| \leq 1} |f(x)|.$$

⊛ **Example 2.2**: For $y \in H$, let $f_y : H \to \mathbb{R}$ be given by $f_y(x) = (x, y)$. By CS,

$$\|f_y\|_{H^*} = \sup_{\|x\| \leq 1} (x, y) \leq \sup_{\|x\| \leq 1} \|x\| \|y\| \leq \|y\|.$$

Also, if $y \neq 0$, then

$$f_y\left(\frac{y}{\|y\|}\right) = \left(\frac{y}{\|y\|}, y\right) = \|y\|.$$

Thus, $\|f_y\|_{H^*} = \|y\|_H$. It turns out all such functionals are of this form.

↪**Theorem 2.8** (Riesz Representation for Hilbert Spaces): If $f \in H^*$, there exists a unique $y \in H$ such that $f(x) = (x, y)$ for every $x \in X$.

Proof. We show first existence. If $f \equiv 0$, then $y = 0$. Otherwise, let $M = \{x \in X \mid f(x) = 0\}$, so $M \subsetneq H$. f linear, so M a linear subspace. f is continuous, so in addition M is closed. By the previous theorem, $M^\perp \neq \{0\}$. Let $z \in M^\perp$ of norm 1.

Fix $x \in H$, and define

$$u := f(x)z - f(z)x.$$

Then, notice that by linearity

$$f(u) = f(x)f(z) - f(z)f(x) = 0,$$

so $u \in M$. Thus, since $z \in M^\perp$, $(u, z) = 0$, so in particular,

$$\begin{aligned}
(u, z) = 0 &= (f(x)z - f(z)x, z) \\
&= f(x)(z, z) - f(z)(x, z) \\
&= f(x)\|z\|^2 - (x, f(z)z) \\
&= f(x) - (x, f(z)z),
\end{aligned}$$

hence, rearranging we find

$$f(x) = (x, f(z)z),$$

and thus letting $y = f(z)z$ completes the proof of existence, noting z independent of x.

For uniqueness, suppose $(x, y) = (x, y')$ for every $x \in X$. Then, $(x, y - y') = 0$ for every $x \in X$, hence letting $x = y - y'$ we conclude $(y - y', y - y') = 0$ thus $y - y' = 0$ so $y = y'$, and uniqness holds. ■

↪**Definition 2.10** (Orthonormal Set): A collection $\{e_j\} \subseteq H$ is *orthonormal* if $(e_i, e_j) = \delta_i^j$.

Remark 2.7: The following section writes notations assuming H has a countable basis. However, for more general Hilbert spaces, all countable summations can be replaced with uncountable ones in which only countably many elements are nonzero. The theory is very similar.

↪**Definition 2.11** (Orthonormal Basis): A collection $\{e_j\} \subseteq H$ is an *orthonormal basis* for H if $\{e_j\}$ is an orthonormal set, and $x = \sum_{j=1}^{\infty} (x, e_j) e_j$ for every $x \in H$, in the sense that

$$\left\| x - \sum_{j=1}^{N} (x, e_j) e_j \right\| \to 0, \qquad N \to \infty.$$

↪**Theorem 2.9** (General Pythagorean Theorem): If $\{e_j\}_{j=1}^{\infty} \subseteq H$ are orthonormal and $\{\alpha_i\}_{i=1}^{\infty} \subseteq \mathbb{R}$ are orthonormal, then for any N,

$$\left\| \sum_{i=1}^{N} \alpha_i e_i \right\|^2 = \sum_{i=1}^{N} |\alpha_i|^2.$$

PROOF.

$$\left\| \sum_{i=1}^{N} \alpha_i e_i \right\|^2 = \left(\sum_{i=1}^{N} \alpha_i e_i, \sum_{j=1}^{N} \alpha_j e_j \right) = \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j \underbrace{(e_i, e_j)}_{=\delta_i^j} = \sum_{i=1}^{N} \alpha_i^2.$$

■

We can also Gram-Schmidt in infinite-dimensional Hilbert spaces. Let $\{x_i\} \subseteq H$. Let

$$e_1 = \frac{x_1}{\|x_1\|},$$

and inductively, for any $n \geq 2$, define

$$v_N = x_N - \sum_{i=1}^{N-1} (x_N, e_i) e_i.$$

Then, for any N, $\operatorname{span}(v_1, ..., v_N) = \operatorname{span}(e_1, ..., e_N)$, and for any $j < N$,

$$(v_N, e_j) = (x_N, e_j) - \sum_{i=1}^{N} (x_N, e_i)(e_i, e_j) = (x_N, e_j) - (x_N, e_j) = 0.$$

Let then $e_N = \frac{v_N}{\|v_N\|}$. Then, $\{e_i\}_{i=1}^{\infty}$ will be orthonormal; we discuss how to establish when this set will actually be a basis to follow.

↪**Theorem 2.10** (Bessel's Inequality): If $\{e_i\}_{i=1}^{\infty}$ are orthonormal, then for any $x \in H$,

$$\sum_{i=1}^{\infty} |(x, e_i)|^2 \leq \|x\|^2.$$

Proof. We have

$$\begin{aligned} 0 &\leq \left\| x - \sum_{i=1}^{N} (x, e_i) e_i \right\|^2 \\ &= \left(x - \sum_{i=1}^{N} (x, e_i) e_i, x - \sum_{j=1}^{N} (x, e_j) e_j \right) \\ &= \|x\| - 2 \sum_{i=1}^{N} (x, e_i)^2 + \sum_{i=1}^{N} (x, e_i)^2 \\ &= \|x\| - \sum_{i=1}^{N} (x, e_i)^2, \end{aligned}$$

so $\sum_{i=1}^{N} (x, e_i)^2 \leq \|x\|$; letting $N \to \infty$ proves the desired inequality, since the RHS is independent of N. ■

↪**Theorem 2.11**: If $\{e_i\}_{i=1}^{\infty}$ are orthonormal, then TFAE:

(a) completeness: if $(x, e_i) = 0$ for every i, then $x = 0$, the zero vector;

(b) Parseval's identity holds: $\|x\|^2 = \sum_{i=1}^{\infty} (x, e_i)^2$ for every $x \in H$;

(c) $\{e_i\}_{i=1}^{\infty}$ form a basis for H, i.e. $x = \sum_{i=1}^{\infty} (x, e_i) e_i$ for every $x \in H$.

Proof. ((a) ⇒ (c)) By Bessel's, $\sum_{i=1}^{\infty} (x, e_i)^2 \leq \|x\|^2$. So, for any $M \geq N$,

$$\left\| \sum_{i=N}^{M} (x, e_i) e_i \right\|^2 = \sum_{i=N}^{M} (x, e_i)^2,$$

which must converge to zero as $N, M \to \infty$, since the whole series converges (being bounded). Hence, $\left\{ \sum_{i=1}^{N} (x, e_i) e_i \right\}_N$ is Cauchy in $\|\cdot\|$ and since H complete, $\sum_{i=1}^{N} (x, e_i) e_i$ converges in H. Putting $y = x - \sum_{i=1}^{\infty} (x, e_i) e_i$, we find

$$(y, e_i) = (x, e_i) - (x, e_i) = 0 \, \forall \, i,$$

hence by assumption in (a), it follows that $y = 0$ so $x = \sum_{i=1}^{\infty} (x, e_i) e_i$ and thus $\{e_i\}$ a basis for H and (c) holds.

((c) ⇒ (b)) Since $x = \sum_{i=1}^{\infty} (x, e_i) e_i$, then,

$$\|x\|^2 - \sum_{i=1}^{N} (x, e_i)^2 = \left\| x - \sum_{i=1}^{N} (x, e_i) e_i \right\|^2 \to 0$$

as $N \to \infty$, hence $\|x\|^2 = \sum_{i=1}^{\infty} (x, e_i)^2$.

((b) $\Rightarrow$ (a)) If $(x, e_i) = 0$ for every i, then by Parseval's $\|x\|^2 = \sum_{i=1}^{\infty} 0 = 0$ so $x = 0$. ■

Remark 2.8: (a) is equivalent to $\mathrm{span}(e_1, e_2, \ldots,)$ is *dense* in H.

↪**Theorem 2.12**: Every Hilbert space has an orthonormal basis.

PROOF. Let $\mathcal{F} = \{\text{orthonormal subsets of } H\}$. $\mathcal{F}$ can be (partially) ordered by inclusion, as can be upper bounded by the union over the whole space. By Zorn's Lemma, there is a maximal set in $\mathcal{F}$, which implies completeness, (a). ■

↪**Proposition 2.5**: H is separable iff H has a countable basis.

PROOF. ($\Leftarrow$) If H has a countable basis $\{e_j\}$, $\mathrm{span}_{\mathbb{Q}}\{e_j\}$ is a countable dense set.

($\Rightarrow$) If H is separable, let $\{x_n\}$ be a countable dense set. Use Gram-Schmidt, to produce a countable, orthonormal set, which is dense and hence a (countable) basis for H. ■

Remark 2.9: All this can be extended to uncountable bases.

§2.6 Adjoints, Duals and Weak Convergence (for Hilbert Spaces)

First consider $T : H \to H$ bounded and linear. Fix $y \in H$. We claim that the map

$$x \mapsto (T(x), y)$$

belongs to H^*, namely is bounded and linear. Linearity is clear since T linear. We know by Cauchy-Schwarz that

$$|(T(x), y)| \le \|T(x)\| \|y\| \le \|T\| \|x\| \|y\| \le C\|x\|,$$

so indeed $x \mapsto (T(x), y) \in H^*$. By Riesz Representation Theorem, there is some unique $z \in H$ such that

$$(T(x), y) = (x, z) \, \forall \, x \in H.$$

This motivates the following.

↪**Definition 2.12** (Adjoint of T): Let $T^* : H \to H$ be defined by

$$(Tx, y) = (x, T^* y), \forall \, x, y \in H.$$

Remark 2.10: In finite dimensions, T can be identified with some $n \times n$ matrix, in which case $T^* = T^t$, the transpose of T; namely $Tx \cdot b = x \cdot T^t b$.

↪**Proposition 2.6**: If $T \in \mathcal{L}(H) := \mathcal{L}(H, H)$, then $T^* \in \mathcal{L}(H)$ and $\|T^*\| = \|T\|$.

Proof. Linearity of T^* is clear. Also, for any $\|y\| \leq 1$,

$$\|T^* y\|^2 = (T^* y, T^* y) = (T T^* y, y) \leq \|T\| \|T^*(y)\| \|y\|$$

so $\|T^* y\| \leq \|T\|$ for all $\|y\| = 1$. so $\|T^*\| \leq \|T\|$ hence $T^* \in \mathcal{L}(H)$. But also, if $x \in H$ with $\|x\| = 1$, then symmetrically,

$$\|Tx\|^2 = (Tx, Tx) = (x, T^* T x) \leq \|T^*\| \|Tx\|$$

so similarly $\|T\| \leq \|T^*\|$ hence equality holds. ■

↪**Proposition 2.7**: $(T^*)^* = T$.

Proof. On the one hand,

$$(T^* y, x) = (y, (T^*)^* x) = ((T^*)^* x, y)$$

while also

$$(T^* y, x) = (x, T^* y) = (Tx, y)$$

so $(Tx, y) = ((T^*)^*, y)$, from which it follows that $T = T^{**}$. ■

↪**Proposition 2.8**: $(T + S)^* = T^* + S^*$, and $(T \circ S)^* = S^* \circ T^*$.

We'll write $N(T)$ for the nullspace/kernel of T, and $R(T)$ for the range/image of T.

↪**Proposition 2.9**: Suppose $T \in \mathcal{L}(H)$. Then,

- $N(T^*) = R(T)^\perp$ (and hence, if $R(T)$ closed, $H = N(T^*) \oplus R(T)$);
- $N(T) = R(T^*)^\perp$ (and hence, if $R(T^*)$ closed, $H = N(T) \oplus R(T^*)$).

Proof. $N(T^*) = \{y \in H : T^* y = 0\}$, so if $y \in N(T^*)$, $(Tx, y) = (x, T^* y) = (x, 0) = 0$, which holds iff y orthogonal to Tx, and since this holds for all $x \in H$, $y \in R(T)^\perp$.

Then, if $R(T)$ closed, the by orthogonal decomposition we'll find $H = R(T) \oplus R(T)^\perp = R(T) \oplus N(T^*)$.

The other claim follows similarly. ■

Remark 2.11: Recall that $R(T)^{\perp}$ is closed; hence

$$\left(R(T)^{\perp}\right)^{\perp} = \left\{z \in H \mid (y, z) = 0 \,\forall\, y \in R(T)^{\perp}\right\},$$

and is also closed; hence $\left(R(T)^{\perp}\right)^{\perp} = \overline{R(T)}$ thus equivalently $N(T^*)^{\perp} = \overline{R(T)}$.

Remark 2.12: By the Closed Graph Theorem, T linear and bounded gives T closed; namely, the graph of T closed; this is *not* the same as saying the range of T closed.

⊛ **Example 2.3**: Consider $C([0,1]) \subseteq L^2([0,1])$, and $T : C([0,1]) \to L^2([0,1])$ given by the identity, $Tf = f$. Then, T is bounded, but $R(T) = C([0,1])$; this subspace is *not* closed in $L^2([0,1])$, since there exists sequences of continuous functions that converge to an L^2, but not continuous, function.

Remark 2.13: The prior theorem is key in "solvability", especially if T a differential or integral operator. If we wish to find u such that $Tu = f$, we need that $f \in R(T)$, hence $f \in N(T^*)^{\perp}$.

⊛ **Example 2.4**: Let $M \subsetneq H$ a closed linear subspace. Then, $H = M \oplus M^{\perp}$; define the projection operator

$$P : H \to H, \qquad x = u + v \in M \oplus M^{\perp} \mapsto u.$$

This means, in particular, $x = Px + (\mathrm{id} - P)x$. We claim $P \in \mathcal{L}(H)$, $\|P\| = 1$, $P^2 = P$, and $P^* = P$.

Linearity is clear. To show $P^2 = P$, write $x = Px + v$. Then, composing both sides with P, we find $Px = P^2x + Pv = P^2x$, so $Px = P^2x$ for every $x \in H$. To see the norm, we find that for every $x \in H$,

$$\begin{aligned}
\|x\|^2 = (x, x) &= (Px + (\mathrm{id} - P)x, Px + (\mathrm{id} - P)x) \\
&= \|Px\|^2 + 2\underbrace{(Px, (\mathrm{id} - P)x)}_{\perp} + \|(\mathrm{id} - P)x\|^2 \\
&= \|Px\|^2 + \|(\mathrm{id} - P)x\|^2 \geq \|Px\|^2 \\
&\Rightarrow \|Px\| \leq \|x\| \Rightarrow \|P\| \leq 1,
\end{aligned}$$

and moreover if $x \in M$, $Px = x$ so $\|Px\| = \|x\|$ hence $\|P\| = 1$ indeed.

Finally, the show P self-adjoint, let $x, y \in H$, then,

$$0 = (Px, (\mathrm{id} - P)y) = (Px, y - Py) \Rightarrow (Px, y) = (Px, Py).$$

Symmetrically, $(x, Py) = (Px, Py)$, hence $(Px, y) = (x, Py)$, and so $P = P^*$.

§2.7 Introduction to Weak Convergence

We let throughout X be a Banach space.

↪**Definition 2.13** (Weak convergence): We say $\{x_n\} \subseteq X$ *converges weakly* to $x \in X$, and write

$$x_n \rightharpoonup x$$

iff for every $f \in X^* = \{f : X \to \mathbb{R} \text{ bounded, linear}\}$, $f(x_n) \to f(x)$.

↪**Definition 2.14** (Weak topology $\sigma(X, X^*)$): The weak topology $\sigma(X, X^*)$ is the weak topology induced by

$$\mathcal{F} = X^*.$$

In particular, this is the smallest topology in which every f continuous.

Recall that this was defined as being $\tau(\{f^{-1}(\mathcal{O})\})$ for $\mathcal{O}$ open in $\mathbb{R}$. A base for this topology is given by $\mathcal{B} = \{\text{finite intersections of } \{f^{-1}\mathcal{O}\}\}$. Namely, let $\mathcal{B}_X := \left\{B_{\varepsilon, f_1, f_2, \ldots, f_n}(x)\right\}$ where

$$B_{\varepsilon, f_1, f_2, \ldots, f_n}(x) = \{x' \in X \mid |f_k(x') - f_k(x)| < \varepsilon, \forall\, 1 \leq k \leq n\}.$$

So, $x_n \to x$ in $\sigma(X, X^*)$ if for every $\varepsilon > 0$, and ball $B_{\varepsilon, f_1, \ldots, f_m}(x)$, there is an N such that for every $n \geq N$, $x_n \in B_{\varepsilon, f_1, \ldots, f_m}(x)$, hence for every $f \in X^*$, $|f(x_n) - f(x)| < \varepsilon$.

For Hilbert spaces, by Riesz we know $f \in H^*$ can always be identified with $f(x) = (x, y)$ for some $y \in H$. So, we find $x_n \rightharpoonup x$ in H iff for every $y \in H$, $(x_n, y) \to (x, y)$.

Remark 2.14: If $x_n \to x$ in H, then $(x_n, y) \to (x, y)$; so this "normal" (we say "strong") convergence implies weak convergence.

↪**Proposition 2.10**: (i) Suppose $x_n \rightharpoonup x$ in H. Then, $\{x_n\}$ are bounded in H, and $\|x\| \leq \liminf_{n \to \infty} \|x_n\|$.

(ii) If $y_n \to y$ (strongly) in H and $x_n \rightharpoonup x$ (weakly) in H, then $(x_n, y_n) \to (x, y)$.

Remark 2.15: It does *not* hold, though, that $x_n \rightharpoonup x$, $y_n \rightharpoonup y$ gives $(x_n, y_n) \to (x, y)$.

Proof. (i) If $x_n \rightharpoonup x$, then

$$\left(x_n, \frac{x}{\|x\|}\right) \to \left(x, \frac{x}{\|x\|}\right) = \|x\|.$$

By Cauchy-Schwarz, we also have

$$\left|\left(x_n, \frac{x}{\|x\|}\right)\right| \leq \|x_n\| \left(\frac{\|x\|}{\|x\|}\right) = \|x_n\|,$$

hence we conclude

$$\liminf_{n\to\infty}\left(x_n, \frac{x}{\|x\|}\right) \le \liminf_{n\to\infty}\|x_n\| \Rightarrow \|x\| \le \liminf_{n\to\infty}\|x_n\|.$$

To argue $\{x_n\}$ bounded, we need the uniform boundedness principle. We can view $\{x_n\} \subseteq H^{**}$ by the canonical association $x_n^{**} : f \mapsto f(x_n)$. Since $f \in H^*$, there is a y such that $f(\cdot) = (\cdot, y)$; label $f = f_y$. Then, for every $f \in H^*$,

$$x_n^{**}(f_y) = f_y(x_n) = (x_n, y) \to (x, y),$$

by weak convergence. Hence, it must be that $\sup_n |x_n^{**} f| = \sup_n |f_y(x_n)| < \infty$ for every $f \in H^*$, namely $\{x_n^{**}\}$ is a pointwise-bounded family of functions. Thus, by uniform boundedness, there is a $C > 0$ such that $|x_n^{**} f| \le C\ \|f\|$ for every $f \in H^*$ and $n \ge 1$. In particular, if we take $f(\cdot) := (\cdot, x_n)$, we know by Riesz that $\|f\| = \|x_n\|$ on the one hand, so for every $n \ge 1$,

$$C\|f\| = C\ \|x_n\| \ge |x_n^{**} f| = |(x_n, x_n)| = \|x_n\|^2 \Rightarrow \|x_n\| \le C,$$

completing the claim of boundedness.

(ii) If $y_n \to y$ in H,

$$\begin{aligned} |(x_n, y_n) - (x, y)| &\le |(x_n, y_n - y)| + |(x_n - x, y)| \\ &\le \underbrace{\|x_n\|}_{\text{bounded}}\ \underbrace{\|y_n - y\|}_{\to 0} + \underbrace{|(x_n - x, y)|}_{\to 0 \text{ by weak}} \to 0. \end{aligned}$$

■

The real help of weak convergence is in the ease of achieving weak compactness;

↪**Theorem 2.13** (Weak Compactness): Every bounded sequence in H has a weakly convergent subsequence.

↪**Theorem 2.14** (Helley's Theorem): Let X a separable normed vector space and $\{f_n\} \subseteq X^*$ such that there is a constant $C > 0$ such that $|f_n(x)| \le C\|x\|$ for every $x \in X$ and $n \ge 1$. Then, there exists a subsequence $\left\{f_{n_k}\right\}$ and an $f \in X^*$ such that $f_{n_k}(x) \to f(x)$ for every $x \in X$.

Proof. This is essentially a specialization of the Arzelà-Ascoli lemma. To apply it, we need X separable (done), the sequence to be pointwise bounded (done), and the sequence to be equicontinuous. To verify this last one, we know that

$$\|f_n(x)\| \le C\|x\| \Rightarrow \|f_n\| \le C, \forall\, n \ge 1,$$

hence by linearity, for any $x, y \in X$,

$$\|f_n(x) - f_n(y)\| \le C\|x - y\|, \forall\, n \ge 1,$$

so in particular $\{f_n\}$ uniformly Lipschitz, thus equicontinuous. ■

PROOF. (Of Thm. 2.13) Let $\{x_n\} \subseteq H$ be bounded and let $H_0 = \overline{\text{span}\{x_1, ..., x_n, ...\}}$, so H_0 is separable, and $(H_0, (\cdot, \cdot))$ is a Hilbert space (being closed). Let $f_n \in H_0^*$ be given by

$$f_n(x) = (x_n, x), \forall\, x \in H_0.$$

Then,

$$|f_n(x)| \leq \|x_n\| \|x\| \leq C\|x\|,$$

since $\{x_n\}$ bounded by assumption. By Helly's Theorem, then, there is a subsequence $\left\{f_{n_k}\right\}$ such that $f_{n_k}(x) \to f(x)$ for every $x \in H_0$, where $f \in H_0^*$. By Riesz, then, $f(x) = (x, x_0)$ for some $x_0 \in H_0^*$. This implies

$$\left(x_{n_k}, x\right) \to (x_0, x), \forall\, x \in H_0.$$

Let P the projection of H onto H_0. Then, for every $x \in H$,

$$\left(x_{n_k}, (\text{id} - P)x\right) = (x_0, (\text{id} - P)x) = 0$$

so for any $x \in H$,

$$\begin{aligned}
\lim_{k\to\infty} \left(x_{n_k}, x\right) &= \lim_{k\to\infty} \left(x_{n_k}, Px + (\text{id} - P)x\right) \\
&= \lim_{k\to\infty} (x_{n_k}, \underbrace{Px}_{\in H_0}) \\
&= (x_0, Px) = (x_0, Px + (\text{id} - P)x) = (x_0, x),
\end{aligned}$$

as we aimed to show. ■

§2.8 Review of L^p Spaces

We always consider $\Omega \subseteq \mathbb{R}^d$.

↪**Definition 2.15** ($L^p(\Omega)$): For $1 \leq p < \infty$, define

$$L^p(\Omega) := \left\{ f : \Omega \to \mathbb{R} \mid f \text{ measurable and } \int_\Omega |f|^p \, \mathrm{d}x < \infty \right\},$$

endowed with the norm

$$\|f\|_{L^p(\Omega)} = \|f\|_p := \left[\int_\Omega |f(x)|^p \, \mathrm{d}x\right]^{\frac{1}{p}}.$$

For $p = \infty$, define

$$L^\infty(\Omega) = \{f : \Omega \to \mathbb{R} \mid f \text{ measurable and} \,\exists\, C < \infty \text{ s.t. } |f| \leq C \text{ a.e.}\},$$

endowed with the norm

$$\|f\|_{L^\infty(\Omega)} = \|f\|_\infty := \inf\{C : |f| \leq C \text{ a.e.}\}.$$

The following are recalled but not proven here, see here.

↪**Theorem 2.15** (Holder's Inequality): For $1 \le p, q \le \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$, then if $f \in L^p(\Omega), g \in L^q(\Omega)$, then $fg \in L^1(\Omega)$, and

$$\int |fg| \, \mathrm{d}x \le \|f\|_p \|g\|_q.$$

↪**Theorem 2.16** (Minkowski's Inequality): For all $1 \le p \le \infty$, $\|f+g\|_p \le \|f\|_p + \|g\|_p$. In particular, $L^p(\Omega)$ is a normed vector space.

↪**Theorem 2.17** (Riesz-Fischer Theorem): $L^p(\Omega)$ is a Banach space for every $1 \le p \le \infty$.

↪**Theorem 2.18**: $C_c(\mathbb{R}^d)$, the space of continuous functions with compact support, simple functions, and step functions are all dense subsets of $L^p(\mathbb{R}^d)$, for every $1 \le p < \infty$.

↪**Theorem 2.19** (Separability of $L^p(\Omega)$): L^p is separable, for every $1 \le p < \infty$.

Proof. We prove for $\Omega = \mathbb{R}^d$. Let

$$\mathcal{R} := \left\{ \prod_{i=1}^{d} (a_i, b_i) \mid a_i, b_i \in \mathbb{Q} \right\},$$

and let

$$\mathcal{E} := \{\text{finite linear combinations of } \chi_R \text{ for } R \in \mathcal{R} \text{ with coefficients in } \mathbb{Q}\},$$

where χ_R the indicator function of the set R. Then, we claim $\mathcal{E}$ dense in $L^p(\mathbb{R}^d)$.

Given $f \in L^p(\mathbb{R}^d)$ and $\varepsilon > 0$, by density of $C_c(\mathbb{R}^d)$ there is some f_1 with $\|f - f_1\|_p < \varepsilon$. Let $\operatorname{supp}(f_1) \subseteq R \in \mathcal{R}$. Now, let $\delta > 0$. Write

$$R = \cup_{i=1}^{N} R_i, \qquad R_i \in \mathcal{R},$$

such that

$$\operatorname{osc}_{R_i}(f_1) := \sup_{R_i} f_1 - \inf_{R_i} f_1 < \delta.$$

Then, let

$$f_2(x) = \sum_{i=1}^{N} q_i \chi(R_i), \qquad q_i \in \mathbb{Q} \text{ s.t. } q_i \approx f_1|_{R_i},$$

so

$$\|f_2 - f_1\|_\infty < \delta.$$

Hence,

$$\|f_2 - f_1\|_p \le \left(\int_R |f_2(x) - f_1(x)|^p \, \mathrm{d}x \right)^{\frac{1}{p}}$$
$$\le |f_1 - f_2|_\infty \cdot m(R)^{\frac{1}{p}} < \delta \cdot m(R)^{\frac{1}{p}},$$

where m the Lebesgue measure on $\mathbb{R}^d$. δ was arbitrary so we may take it arbitrarily small such that $\delta m(R)^{\frac{1}{p}} < \varepsilon$, hence for such a δ,

$$\|f - f_2\|_p \le \|f - f_1\|_p + \|f_1 - f_2\|_p < 2\varepsilon.$$

Now, $f_2 \in \mathcal{E}$, and thus $\mathcal{E}$ is dense in $L^p(\mathbb{R}^d)$, and countable by construction, thus $L^p(\mathbb{R}^d)$ separable. ■

Remark 2.16: $L^\infty(\Omega)$ is *not* separable, and $C_c(\mathbb{R}^d)$ is *not* dense in $L^\infty(\Omega)$.

Remark 2.17 (Special Cases):
- If Ω has finite measure, $L^p(\Omega) \subseteq L^{p'}(\Omega)$ for every $p \ge p'$.
- $\ell^p := \left\{ a = (a_n)_{n=1}^\infty \mid \sum_{n=1}^\infty |a_n|^p < \infty \right\}$ endowed with the norm $|a|_{\ell^p} := \left(\sum_{n=1}^\infty |a_n|^p \right)^{1/p}$.

§2.9 $(L^p)^*$: The Riesz Representation Theorem

We are interested in functions $T : L^p(\Omega) \to \mathbb{R}$ which is bounded and linear. For instance, let $g \in L^q(\Omega)$ and $f \in L^p(\Omega)$ where p, q conjugates, and define

$$T(f) := \int_\Omega f(x) g(x) \, \mathrm{d}x.$$

This is clearly linear, and by Holders,

$$|Tf| = \left| \int_\Omega fg \right| \le \|f\|_p \|g\|_q.$$

so

$$\left| T\left(\frac{f}{\|f\|_p} \right) \right| \le \|g\|_q, \ \forall f \in L^p(\Omega), \Rightarrow \|T\| \le \|g\|_q,$$

and thus $T \in (L^p(\Omega))^*$. Moreover, if $1 < p < \infty$, $1 < q < \infty$, let

$$f(x) = \frac{|g(x)|^{q-2} g(x)}{\|g\|_q^{q-1}}.$$

Then,

$$\begin{aligned}\int_\Omega |f(x)|^p \, \mathrm{d}x &= \frac{1}{\|g\|_q^{(q-1)p}} \int_\Omega |g(x)|^{(q-2)p} |g(x)|^p \, \mathrm{d}x \\ &= \frac{1}{\|g\|_q^{(q-1)p}} \int_\Omega |g(x)|^{qp-p} \, \mathrm{d}x.\end{aligned}$$

Since $\frac{1}{p} + \frac{1}{q} = 1$, we have $q + p = pq$, so further

$$= \frac{1}{\|g\|_q^q} \int_\Omega |g(x)|^q \, \mathrm{d}x = \frac{1}{\|g\|_q^q} \cdot \|g\|_q^q = 1,$$

so f as defined indeed in $L^p(\Omega)$ and moreover has L^p-norm of 1. In addition,

$$\begin{aligned}|Tf| &= \frac{1}{\|g\|_q^{q-1}} \int_\Omega \left|g(x)^{q-2}\right| g(x) g(x) \, \mathrm{d}x \\ &= \frac{1}{\|g\|_q^{q-1}} \int_\Omega |g(x)|^q \, \mathrm{d}x \\ &= \frac{1}{\|g\|_q^{q-1}} \|g\|_q^q = \|g\|_q,\end{aligned}$$

so $\|T\| = \|g\|_q$ as desired. We have, more generally, akin to the Riesz representation theorem,

↪**Theorem 2.20** (Riesz-Representation Theorem for $L^p(\Omega)$): Let $1 \leq p < \infty$. For any $T \in (L^p(\Omega))^*$, there exists a unique $g \in L^q(\Omega)$ such that $T(f) = \int_\Omega f(x) g(x) \, \mathrm{d}x$ with $\|T\| = \|g\|_q$.

We'll only prove for $\Omega \subseteq \mathbb{R}$. First:

↪**Proposition 2.11**: Let $T, S \in (L^p(\Omega))^*$. If $T = S$ on a dense subset $E \subseteq L^p(\Omega)$, then $T = S$ everywhere.

PROOF. Let $f_0 \in L^p(\Omega)$. By density, there exists $\{f_n\} \subseteq E$ such that $f_n \to f$ in $L^p(\Omega)$. By continuity, $Tf_n \to Tf_0$ and $Sf_n \to Sf_0$, while $Tf_n = Sf_n$ for every $n \geq 1$, so by uniqueness of limits in $\mathbb{R}$, $Tf_0 = Sf_0$. ■

The general outline of the proof of Thm. 2.20 is the following:

- prove the theorem for f a step function;
- prove the theorem for f bounded and measurable;
- conclude the full theorem by appealing to the previous proposition.

To do this, we need first to recall the notion of *absolutely continuous functions*. Fix $[a, b] \subseteq R$ and $G : [a, b] \to \mathbb{R}$. G is said to be absolutely continuous on $[a, b]$ if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for every disjoint collection $\{(a_k, b_k)\}_{k=1}^N \subseteq [a, b]$ with $\sum_{k=1}^N (a_k - b_k) < \delta$, then $\sum_{k=1}^N |G(b_k) - G(a_k)| < \varepsilon$. In particular, we need the following result, proven here:

↪**Theorem 2.21**: If $G : [a, b] \to \mathbb{R}$ is absolutely continuous, then $g = G'$ exists a.e. on $[a, b]$, $g \in L^1([a, b])$, and for every $x \in [a, b]$,

$$G(x) - G(a) = \int_a^x g(t) \,\mathrm{d}t.$$

Proof (*Of Thm. 2.20 with* $\Omega = [a, b]$). Let $T \in (L^p([a, b]))^*$.

Step 1: Let f a step function. The function $\chi_{[a,x)} \in L^p([a, b])$; define

$$G_T(x) := T\big(\chi_{[a,x)}\big).$$

We claim G_T absolutely continuous. Consider $\{(a_k, b_k)\}_{k=1}^N$ disjoint. Then, for every $[c, d] \subseteq [a, b]$, , $G_T(d) - G_T(c) = T\big(\chi_{[a,d)}\big) - T\big(\chi_{[a,c]}\big) = T\big(\chi_{[a,d)} - \chi_{[a,c)}\big) = T\big(\chi_{[c,d)}\big)$, so

$$\begin{aligned}
\sum_{k=1}^N (G_T(b_k) - G_T(a_k)) &= \sum_{k=1}^N c_k \cdot (G_T(b_k) - G_T(a_k)), \qquad c_k := \operatorname{sgn}(G_T(b_k) - G_T(a_k)) \\
&= \sum_{k=1}^N c_k \cdot T\big(\chi_{[a_k,b_k)}\big) \\
&= T\left(\sum_{k=1}^N c_k \chi_{[a_k,b_k)}\right) \\
&\leq \|T\| \left\| \sum_{k=1}^N c_k \chi_{[a_k,b_k)} \right\|_p .
\end{aligned}$$

By the disjointedness of the intervals, we may write

$$\int_a^b \left| \sum_{k=1}^N c_k \chi_{[a_k,b_k)} \right|^p \mathrm{d}x \leq \sum_{k=1}^N \int_{a_k}^{b_k} \mathrm{d}x = \sum_{k=1}^N (b_k - a_k).$$

So, $\left\|\sum_{k=1}^N c_k \chi_{[a_k,b_k)}\right\|_p = \left(\sum_{k=1}^N (b_k - a_k)\right)^{\frac{1}{p}}$, thus

$$\sum_{k=1}^N |G_T(b_k) - G_T(a_k)| \leq \|T\| \cdot \left(\sum_{k=1}^N (b_k - a_k)\right)^{\frac{1}{p}}.$$

Hence, for $\varepsilon > 0$, letting $\delta = \left(\frac{\varepsilon}{\|T\|}\right)^p$ proves absolute continuity of G_T. Thus, $g = G_T'$ exists and is such that $g \in L^1([a, b])$ and

$$G_T(x) = \int_a^x g(t) \,\mathrm{d}t, \ \forall\, x \in [a, b].$$

Hence,

$$
\begin{aligned}
T\Big(\chi_{[c,d)}\Big) = G_T(d) - G_T(c) &= \int_a^d g(t)\,\mathrm{d}t - \int_a^c g(t)\,\mathrm{d}t \\
&= \int_c^d g(t)\,\mathrm{d}t \\
&= \int_a^b g(t)\cdot \chi_{[c,d)}(t)\,\mathrm{d}t.
\end{aligned}
$$

This proves the theorem for indicator functions; by linearity of T and linearity of the integral, we can repeat this procedure to find a function g such that $Tf = \int_a^b f(t)g(t)\,\mathrm{d}t$ for every step function f.

Step 2: Let f bounded and measurable. We know that for every step function ψ, $T\psi = \int_a^b \psi(t)g(t)\,\mathrm{d}t$ (with the g as "found" in step 1). So,

$$
\begin{aligned}
\left|Tf - \int_a^b f(t)g(t)\right| &= \left|T(f-\psi) - \int_a^b (f(t)-\psi(t))g(t)\,\mathrm{d}t\right| \\
&\leq \|T\|\|f-\psi\|_p + \int_a^b |f(t)-\psi(t)||g(t)|\,\mathrm{d}t.
\end{aligned}
$$

Then, since $g \in L^1([a,b])$, for every $\varepsilon > 0$ there is some $\delta > 0$ such that if E a set of measure less than δ, $\int_E |g(t)|\,\mathrm{d}t < \varepsilon$. Fix $\varepsilon > 0$ and $\delta > 0$ such that this holds; let $\delta < \varepsilon$ if necessary wlog. Since f bounded and measurable, there is some step function ψ such that $|f-\psi| < \delta$ on $E \subseteq [a,b]$, and that $m(E^c) < \delta$ and $|\psi| \leq \|f\|_\infty$. Hence,

$$
\begin{aligned}
\|f-\psi\|_p^p &= \int_E |f-\psi|^p + \int_{E^c} |f-\psi|^p \\
&\leq \delta^p \cdot m(E) + \Big(2\|f\|_\infty\Big)^p m(E^c) \\
&\leq \delta^p\ |b-a| + \Big(2\|f\|_\infty\Big)^p \delta.
\end{aligned}
$$

Also,

$$
\begin{aligned}
\int_a^b |f-\psi||g|\,\mathrm{d}t &\leq \int_E \delta\cdot|g|\,\mathrm{d}t + \int_{E^c} 2\|f\|_\infty |g|\,\mathrm{d}t \\
&\leq \delta\|g\|_1 + 2\|f\|_\infty \varepsilon.
\end{aligned}
$$

All together then,

$$
\begin{aligned}
\left|Tf - \int_a^b f(t)g(t)\,\mathrm{d}t\right| &\leq \|T\|\Big(\delta^p\ |b-a| + \Big(2\|f\|_\infty\Big)^p\delta\Big)^{\frac{1}{p}} + \delta\|g\|_1 + 2\|f\|_\infty\varepsilon \\
&< C\Big(\|f\|_\infty, \|g\|_1, a, b, \|T\|\Big)\cdot \varepsilon^{\frac{1}{p}},
\end{aligned}
$$

where C a constant. The LHS does not depend on ε, hence taking the limit $\varepsilon \to 0^+$, we conclude

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

Note that all simple functions are bounded and measurable, so the necessary property also holds for f simple.

We need now to show $g \in L^q([a,b])$ and $\|g\| = \|T\|$.

- Case 1: $p > 1$ so $q < \infty$. Let $g_n := \begin{cases} g & \text{if } |g| \le n \\ 0 & \text{o.w.} \end{cases}$ and $f_n := \begin{cases} |g|^{q-1}\,\mathrm{sgn}(g) & \text{if } |g| \le n \\ 0 & \text{o.w.} \end{cases}$. Then,

$$\begin{aligned} \|g_n\|_q^q &= \int_{\{|g|\le n\}} |g|^q\,\mathrm{d}t \\ &= \int_{\{|g|\le n\}} f_n \cdot g_n\,\mathrm{d}t \\ &= \int_{\{|g|\le n\}} f_n g\,\mathrm{d}t \\ &= Tf_n \le \|T\|\|f_n\|_p, \end{aligned}$$

  since f_n bounded and measurable so Step 2 applies. Also,

$$\begin{aligned} \|f_n\|_p^p &= \int_{\{|g|\le n\}} |g|^{(q-1)p}\,\mathrm{d}t \\ &= \int_{\{|g|\le n\}} |g|^q\,\mathrm{d}t = \|g_n\|_q^q. \end{aligned}$$

  All together then,

$$\|g_n\|_q^q \le \|T\|\|g_n\|_q^{q/p} \Rightarrow \|g_n\|_q^{q\left(1-\frac{1}{p}\right)} = \|g_n\|_q \le \|T\|.$$

  By construction, $|g_n|^q \to |g|^q$ a.e. and monotonely, so by the monotone convergence theorem,

$$\|g_n\|_q \to \|g\|_q,$$

  so $\|g\|_q \le \|T\|$ and so $g \in L^q([a,b])$. From here, as in the example at the beginning of this section, one can show equality by chosing f appropriately.

- Case 2: $p = 1$ so $q = \infty$. We claim that $\|g\|_\infty = \sup_{\substack{\|f\|_1=1, \\ f \text{ bdd}}} \int fg$. Let $\varepsilon > 0$ and $A \subseteq [a,b]$ such that $|g| \ge \|g\|_\infty - \varepsilon$ on A where $m(A) > 0$. Let

$$f(x) = \frac{\chi_A}{m(A)}\,\mathrm{sgn}(g).$$

  Then, f bounded and $\|f\|_1 = 1$. So,

$$\int fg = \frac{1}{m(A)} \int_A |g| \ge \frac{1}{m(A)} \int_A \left(\|g\|_\infty - \varepsilon\right) = \|g\|_\infty - \varepsilon,$$

  hence we have proven $\le$ of our claim. By Holder,

$$\sup_{\|f\|=1} \int fg \leq \|f\|_1 \|g\|_\infty = \|g\|_\infty,$$

so $\geq$ holds and the claim is proven. Thus,

$$\|g\|_\infty = \sup_{\substack{\|f\|=1,\\ f \text{ bdd}}} Tf \leq \|T\| \|f\|_1 = \|T\|,$$

so in particular $g \in L^\infty([a, b])$. For the other inequality,

$$|Tf| = \left| \int fg \, \mathrm{d}t \right| \leq \|f\|_1 \|g\|_\infty,$$

hence

$$\|T\| \leq \|g\|_\infty$$

so $\|g\|_\infty = \|T\|$ as we aimed to show.

Step 3: We need to show $Tf = \int_a^b fg \, \mathrm{d}t$ for every $f \in L^p([a, b])$. Simple functions are dense in $L^p([a, b])$, and since $Tf = \int_a^b fg \, \mathrm{d}t$ for every simple function f, we conclude $Tf = \int_a^b fg \, \mathrm{d}t$ for every $f \in L^p([a, b])$ by the previous density lemma.

Moreover, g is unique because if

$$\int_a^b fg = \int_a^b fg',$$

then

$$\int_a^b f(g - g') = 0,$$

for every $f \in L^p$. Let $f(t) = \operatorname{sgn}(g - g')$, then

$$0 = \int_a^b |g - g'| \, \mathrm{d}t \Rightarrow g = g' \text{ a.e..}$$

So, g uniquely defined up to a set of measure 0 so $g = g'$ in L^q. ■

Proof *(Of RRT if $\Omega = \mathbb{R}$)*. Fix $T \in (L^p(\mathbb{R}))^*$. Then, $T|_{[-N,N]} \in (L^p([-N, N]))^*$ for every $N \geq 1$, and $\left\|T|_{[-N,N]}\right\| \leq \|T\|$. Then, by RRT on $[-N, N]$, there is a $g_N \in L^q([-N, N])$ such that $Tf = \int_{-N}^N fg_N \, \mathrm{d}t$. By uniqueness, $g_{N+1}|_{[-N,N]} = g_N$. Define

$$g(t) := g_N(t), \qquad t \in [-N, N].$$

So, $g_N(t) \to g(t)$ pointwise and $|g_N(t)|^q \to |g(t)|^q$ pointwise and monotonely. By monotone convergence, then, $\int_\mathbb{R} |g_N|^q \, \mathrm{d}t \to \int_\mathbb{R} |g|^q \, \mathrm{d}t$. So, $g \in L^q(\mathbb{R})$ since $\|g_N\|_{L^q([-N,N])} \leq \|T\|$ for every $N \geq 1$. Let $f_N(t) = f(t)\chi_{[-N,N]}$. Then, $f_N \to f$ in $L^p(\mathbb{R})$ so $Tf_N \to Tf$. So also

$$Tf_N = \int_{-N}^N f_N g_N = \int_{-N}^N f(t) g_N(t) \, \mathrm{d}t = \int_\mathbb{R} fg_N \, \mathrm{d}t \to Tf,$$

if we take by convention the g_N's to be zero outside of $[-N, N]$. But also, $f \in L^p(\mathbb{R})$ and $g_N \to g$ in $L^q(\mathbb{R})$, so applying Holder's to the quantity $\int_{\mathbb{R}} f g_N$, we know

$$\int_{\mathbb{R}} f g_N \to \int_{\mathbb{R}} f g,$$

hence equating the two

$$Tf = \int_{\mathbb{R}} fg,$$

for every $f \in L^p(\mathbb{R})$. A similar proof to the previous gives the necessary norm identity. ■

Proof *(Of RRT for general $\Omega \subseteq \mathbb{R}$)*. If $T \in (L^p(\Omega))^*$, let $\hat{T} \in (L^p(\mathbb{R}))^*$ given by $\hat{T}f = T(f\mid_\Omega)$. Then by the previous case there is $\hat{g} \in L^q(\mathbb{R})$ such that $\hat{T}(f) = \int f\hat{g}$. Let $g = \hat{g}|_\Omega$, then $Tf = \int_\Omega fg$. ■

So, RRT gives us that for $p \in [1, \infty]$, $(L^p(\Omega))^* \sim L^q(\Omega)$, and that $\|f\|_p = \sup_{\substack{g \in L^q \\ \|g\|_q = 1}} \left|\int fg\right|$.

In particular, if $p = 1$,

$$\|f\|_{L^1} = \int f \operatorname{sgn} f(x) \,\mathrm{d}x = \sup_{\|g\|_\infty = 1} \int fg.$$

What, though, is $(L^\infty)^*$. Certainly, $L^1(\Omega) \subseteq (L^\infty(\Omega))^*$ since for $f \in L^\infty$, $Tf = \int fg \,\mathrm{d}x$ with $g \in L^1$, which is bounded by Holders. However, it turns out that this inclusion is a strict one. Consider for instance

$$Tf := f(0), \qquad T : L^\infty([-1, 1]) \to \mathbb{R}.$$

Then, certainly $|Tf| \leq \|f\|_\infty$ so $T \in (L^\infty)^*$. However, there is no function g such that $f(0) = \int f(t)g(t)\,\mathrm{d}t$.

§2.10 Weak Convergence in $L^p(\Omega)$

↪**Definition 2.16** (Weak convergence in $L^p(\Omega)$): Let $\Omega \subset \mathbb{R}^d$, $p \in [1, \infty)$ and q its conjugate. Then, we say $f_n \to f$ *weakly* in $L^p(\Omega)$, and write

$$f_n \underset{L^p(\Omega)}{\rightharpoonup} f,$$

if for every $g \in L^q(\Omega)$,

$$\lim_{n \to \infty} \int_\Omega f_n g \,\mathrm{d}x = \int fg \,\mathrm{d}x.$$

Remark 2.18: Weak limits are unique; suppose otherwise that $f_n \rightharpoonup f, \overline{f}$. Let $g = \operatorname{sgn}\left(f - \overline{f}\right) \cdot \left|f - \overline{f}\right|^{p-1}$, which is in $L^q(\Omega)$. So,

$$\lim_n \int g f_n \,\mathrm{d}x = \int g f \,\mathrm{d}x = \int g \overline{f} \,\mathrm{d}x,$$

by assumption, so

$$0 = \int_\Omega g\left(f - \overline{f}\right) \mathrm{d}x = \int \left|f - \overline{f}\right|^p \mathrm{d}x,$$

hence $f = \overline{f}$ a.e. (and so equal as elements of $L^p(\Omega)$).

Remark 2.19: Many of the properties of weakly convergent sequences in a Hilbert space carry over to this setting.

↪**Proposition 2.12**: Let $\Omega \subseteq \mathbb{R}^d$.

(i) If $p \in (1, \infty)$, $f_n \underset{L^p(\Omega)}{\rightharpoonup} f$, then $\{f_n\} \subseteq L^p(\Omega)$ are bounded, and moreover $\|f\|_p \leq \liminf_n \|f_n\|_p$.

(ii) If $p \in [1, \infty)$ and $f_n \underset{L^p(\Omega)}{\rightharpoonup} f, g_n \underset{L^p(\Omega)}{\rightarrow} g$, then $\lim_{n\to\infty} \int g_n f_n \,\mathrm{d}x = \int g f \,\mathrm{d}x$.

Proof. Identical to Hilbert space proofs; replace usage of Cauchy-Schwarz with Holder's. ■

Remark 2.20: In (i), $p \in (1, \infty)$, since L^p "reflexive" in this case, i.e. $(L^p)^{**} = L^p$ (just as we had in the Hilbert space case). We don't have this property for $p = 1$.

Remark 2.21: A related notion of convergence is called *weak* convergence*, written $f_n \overset{*}{\underset{L^p(\Omega)}{\rightharpoonup}} f$; we say this holds if for every $g \in L^q(\Omega)$ such that $(L^q)^* = L^p$, then $\int f_n g \,\mathrm{d}x \to \int f g \,\mathrm{d}x$. So if $p \in (1, \infty)$, weak convergence = weak* convergence, by Riesz.

Remark 2.22: There are many equivalent notions to weak convergence.

↪**Theorem 2.22** (Equivalent Weak Convergence): Let $p \in (1, \infty)$. Suppose $\{f_n\} \subseteq L^p(\Omega)$ are bounded and $f \in L^p$. Then, $f_n \underset{L^p(\Omega)}{\rightharpoonup} f$ iff

- for any $g \in G \subseteq L^q(\Omega)$ such that $\overline{\operatorname{span}(G)} = L^q(\Omega)$, then $\lim_{n\to\infty} \int f_n g = \int f g$;
- $\forall\, A \subseteq \Omega$ measurable with finite measure, then $\lim_{n\to\infty} \int_A f_n \,\mathrm{d}x = \int_A f \,\mathrm{d}x$;
- if $d = 1$ and $\Omega = [a, b]$, then $\lim_{n\to\infty} \int_a^x f_n \,\mathrm{d}x = \int_a^x f \,\mathrm{d}x$ for every $x \in [a, b]$.
- $f_n \to f$ pointwise a.e..

Remark 2.23: Some of these notions extend to $p = 1$, but we state in the $p > 1$ case for simplicity.

↪**Theorem 2.23** (Radon-Riesz): Let $p \in (1, \infty)$. Suppose $f_n \underset{L^p(\Omega)}{\rightharpoonup} f$, then $f_n \underset{L^p(\Omega)}{\to} f$ iff $\lim_{n \to \infty} \|f_n\|_p = \|f\|_p$.

Alternatively, there exists a subsequence $\{f_{n_k}\}$ such that $f_{n_k} \to f$ in $L^p(\Omega)$ iff $\liminf_{n \to \infty} \|f_n\|_p = \|f\|_p$.

PROOF. ($\Rightarrow$) If $f_n \underset{L^p(\Omega)}{\to} f$ then $\|f_n\|_p \to \|f\|_p$ by triangle inequality.

The converse, ($\Leftarrow$), is hard. ■

↪**Theorem 2.24** (Weak Compactness): Let $p \in (1, \infty)$, then every bounded sequence in $L^p(\Omega)$ has a weakly convergent subsequence, with limit in $L^p(\Omega)$.

PROOF. Let $\{f_n\} \subseteq L^p(\Omega)$ be bounded. $p \in (1, \infty)$ so so is q, and in particular $L^q(\Omega)$ is separable. Let $T_n \in (L^q(\Omega))^*$ be given by $T_n(g) := \int f_n g \, \mathrm{d}x$ for $g \in L^q(\Omega)$. Then, $\|T_n\| = \|f_n\|_p \leq C$. So,

$$\sup_n |T_n(g)| \leq \|T_n\| \|g\|_q \leq C \|g\|_q.$$

By Helley's Theorem (Thm. 2.14), there exists a subsequence $\{T_{n_k}\}$ and $T \subseteq (L^q(\Omega))^*$ such that $\lim_{k \to \infty} T_{n_k}(g) = T(g)$ for every $g \in L^q(\Omega)$. By Riesz, there exists some $f \in L^p(\Omega)$ such that $T(g) = \int f g \, \mathrm{d}x$, and hence

$$\lim_k \int f_{n_k} g \, \mathrm{d}x = \int f g \, \mathrm{d}x,$$

for every $g \in L^q(\Omega)$, so $f_{n_k} \underset{L^p(\Omega)}{\rightharpoonup} f$. ■

§2.11 Convolution and Mollifiers

↪**Definition 2.17** (Convolution):

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

↪**Proposition 2.13** (Properties of Convolution):
a. $(f * g) * h = f * (g * h)$ (convolution is associative)
b. Let $\tau_z f(x) := f(x - z)$ be the z-translate of x which centers f at z. Then,

$$\tau_z(f * g) = (\tau_z f) * g = f * (\tau_z g).$$

c. $\operatorname{supp}(f * g) \subseteq \overline{\{x + y \mid x \in \operatorname{supp}(f), y \in \operatorname{supp}(g)\}}$.

PROOF. (a) Assuming all the necessary integrals are finite, we can change order of integration,

$$\begin{aligned}
((f * g) * h)(x) &= \left(\int f(y) g(x - y) \, \mathrm{d}y \right) * h(x) \\
&= \int \int f(y) g(x - z - y) \, \mathrm{d}y, h(z) \, \mathrm{d}z \\
&= \int \int f(y) g(x - y - z) h(z) \, \mathrm{d}z \, \mathrm{d}y \qquad (y' = x - y) \\
&= \int \int f(x - y') g(y' - z) h(z) \, \mathrm{d}z \, \mathrm{d}y' \\
&= \int f(x - y')(g * h)(y') \, \mathrm{d}y' = (f * (g * h))(x).
\end{aligned}$$

(b) For the first equality,

$$\begin{aligned}
\tau_z(f * g)(x) &= \tau_z \int f(x - y) g(y) \, \mathrm{d}y \\
&= \int f(x - z - y) g(y) \, \mathrm{d}y \\
&= \int (\tau_z f(x - y)) g(y) \, \mathrm{d}y = ((\tau_z f) * g)(x).
\end{aligned}$$

The second follows from a change of variables in the second line.

(c) We'll show that $A^c \subseteq (\operatorname{supp}(f * g))^c$ where A the set as defined in the proposition. Let $x \in A^c$, then if $y \in \operatorname{supp}(g)$, $x - y \notin \operatorname{supp}(f)$ so $f(x - y) = 0$; else if $y \notin \operatorname{supp}(g)$ it must be $g(y) = 0$. So, if $x \in A^c$, it must be that

$$\int f(x - y) g(y) \, \mathrm{d}y = \int_{\operatorname{supp}(g)} \underbrace{f(x - y)}_{=0} g(y) \, \mathrm{d}y + \int_{\operatorname{supp}(g)^c} f(x - y) \underbrace{g(y)}_{=0} \, \mathrm{d}y = 0.$$

■

We've been rather loose with finiteness of the convolutions so far. To establish this, we need the following result.

↪**Theorem 2.25** (Young's Inequality): Let $f \in L^1(\mathbb{R}^d), g \in L^p(\mathbb{R}^d)$ for any $p \in [1, \infty]$. Then,

$$\|f * g\|_p \leq \|f\|_1 \|g\|_p,$$

hence $f * g \in L^p(\mathbb{R}^d)$.

PROOF. Suppose first $p = \infty$, then

$$(f * g)(x) = \int f(y)g(x-y)\,\mathrm{d}y \leq \|g\|_\infty \int |f(y)|\,\mathrm{d}y = \|g\|_\infty \|f\|_1,$$

for every $x \in \mathbb{R}^d$, so passing to the L^∞-norm,

$$\|f * g\|_\infty \leq \|f\|_1 \|g\|_\infty.$$

Suppose now $p = 1$. Then,

$$\|f * g\|_1 = \int \left| \int f(x-y)g(y)\,\mathrm{d}y \right| \mathrm{d}x.$$

Let $F(x,y) = f(x-y)g(y)$, then for almost every $y \in \mathbb{R}^d$,

$$\begin{aligned}\int |F(x,y)|\,\mathrm{d}x &= \int |g(y)||f(x-y)|\,\mathrm{d}x \\ &= |g(y)| \int |f(x-y)|\,\mathrm{d}x \\ &= |g(y)|\|f\|_1.\end{aligned}$$

Applying Tonelli's Theorem, we have then

$$\iint |F(x,y)|\,\mathrm{d}y\,\mathrm{d}x = \iint |F(x,y)|\,\mathrm{d}x\,\mathrm{d}y = \int |g(y)|\|f\|_1\,\mathrm{d}y = \|f\|_1 \|g\|_1,$$

(so really $F \in L^1(\mathbb{R}^d) \times L^1(\mathbb{R}^d)$), hence all together

$$\|f * g\|_1 = \int \left| \int F(x,y)\,\mathrm{d}y \right| \mathrm{d}x \leq \iint |F(x,y)|\,\mathrm{d}y\,\mathrm{d}x = \|f\|_1 \|g\|_1.$$

Remark 2.24: It also follows that for a.e. $x \in \mathbb{R}^d$, $\int |F(x,y)|\,\mathrm{d}y < \infty$, i.e. $\int |f(x-y)g(y)|\,\mathrm{d}y < \infty$. Moreover, since if $g \in L^p(\Omega)$ then $|g|^p \in L^1(\Omega)$, a similar argument gives that for almost every $x \in \mathbb{R}^d$, $\int |f(x-y)||g(y)|^p\,\mathrm{d}y < \infty$.

Suppose now $1 < p < \infty$. For a.e. $x \in \mathbb{R}^d$, $\int |g(y)|^p |f(x-y)|\,\mathrm{d}y < \infty$, so $g \in L^p(\mathbb{R}^d)$ implies for a.e. $x \in \mathbb{R}^d$, $|g(\cdot)|^p\ |f(x-\cdot)| \in L^1(\mathbb{R}^d)$ as a function of $\cdot$. This further implies $g(y)f^{\frac{1}{p}}(x-y) \in L^p(\mathbb{R}^d, dy)$. Also, if $f \in L^1(\mathbb{R}^d)$, then $f^{\frac{1}{q}} \in L^q(\mathbb{R}^d)$. All together then,

$$\int |f(x-y)||g(y)|\,\mathrm{d}y = \int \overbrace{\left|f^{\frac{1}{q}}(x-y)\right|}^{q} \underbrace{\left|f^{\frac{1}{p}}(x-y)\right||g(y)|\,\mathrm{d}y}_{p}$$

$$\text{Holder's} \qquad \leq \left(\int |f(x-y)|\,\mathrm{d}y\right)^{\frac{1}{q}} \left(\int |f(x-y)||g(y)|^p\,\mathrm{d}y\right)^{\frac{1}{p}},$$

hence, raising both sides to the p,

$$|(f*g)(x)|^p \leq \|f\|_1^{\frac{p}{q}} \cdot (|f| * |g|^p)(x)$$

and integrating both sides

$$\int |(f*g)(x)|^p\,\mathrm{d}x \leq \|f\|_1^{\frac{p}{q}} \int \left(\underbrace{|f|}_{\in L^1(\mathbb{R}^d)} * \underbrace{|g|^p}_{\in L^1(\mathbb{R}^d)}\right)(x)\,\mathrm{d}x.$$

Hence, we can bound the right-hand term using the previous case for $p = 1$, and find

$$\begin{aligned}
\int |(f*g)(x)|^p\,\mathrm{d}x &\leq \|f\|_1^{\frac{p}{q}} \|f\|_1 \|g^p\|_1 \\
&= \|f\|_1^{\frac{p}{q}+1} \|g\|_p^p \\
&= \|f\|_1^{\frac{p+q}{q}} \|g\|_p^p \\
\left(\frac{p+q}{q} = p\right) \qquad &= \|f\|_1^p \|g\|_p^p,
\end{aligned}$$

so raising both sides to $\frac{1}{p}$, we conclude

$$\|f*g\|_p \leq \|f\|_1 \|g\|_p.$$

■

↪**Proposition 2.14**: If $f \in L^1(\mathbb{R}^d)$ and $g \in C^1(\mathbb{R}^d)$ with $\left|\partial_{x_i} g\right| \in L^\infty(\mathbb{R}^d)$ for $i = 1, ..., d$, then $(f*g) \in C^1(\mathbb{R}^d)$ and moreover

$$\partial_{x_i}(f*g) = f * \left(\partial_{x_i} g\right).$$

Remark 2.25: There are many different conditions we can place on f, g to make this true; most basically, we need $|(\partial_i g) * f| < \infty$.

Proof.

$$\frac{\partial}{\partial x_i}\left(\int f(y) g(x-y)\,\mathrm{d}y\right) = \int \underbrace{f(y)}_{\in L^1(\mathbb{R}^d)} \underbrace{\partial_i g(x-y)}_{\in L^\infty(\mathbb{R}^d)}\,\mathrm{d}y < \infty,$$

citing the previous theorem for the finiteness; the dominated convergence theorem allows us to pass the derivative inside. ■

Remark 2.26: This also follows for the gradient; namely $\nabla(f * g) = f * (\nabla g)$ with a component-wise convolution.

Consider the function

$$\rho(x) = \begin{cases} C \exp\left(-\frac{1}{1-|x|^2}\right) & \text{if } |x| \leq 1 \\ 0 & \text{o.w.} \end{cases},$$

where $C = C(d)$ a constant such that $\int_{\mathbb{R}^d} \rho(x) \, \mathrm{d}x = 1$. Then, note that $\rho \in C_c^\infty(\mathbb{R}^d)$ (infinitely differentiable with compact support). Let now

$$\rho_\varepsilon(x) := \frac{1}{\varepsilon^d} \rho\left(\frac{x}{\varepsilon}\right).$$

Notice that $\rho_\varepsilon(x)$ is supported on $B(0, \varepsilon)$, but

$$\int_{\mathbb{R}^d} \rho_\varepsilon(x) \, \mathrm{d}x = \frac{1}{\varepsilon^d} \int_{\mathbb{R}^d} \rho\left(\frac{x}{\varepsilon}\right) \mathrm{d}x = \frac{1}{\varepsilon^d} \cdot \varepsilon^d \cdot \int_{\mathbb{R}^d} \rho(y) \, \mathrm{d}y = 1,$$

for every ε, by making a change of variables $y = \frac{x}{\varepsilon}$. We'll be interested in the convolution

$$f_\varepsilon(x) := (\rho_\varepsilon * f)(x)$$

for some function f. ρ_ε is often called a "convolution kernel". In particular, it is a "good kernel", namely has the properties:

- $\int_{\mathbb{R}^d} \rho_\varepsilon(y) \, \mathrm{d}y = 1$;
- $\int_{\mathbb{R}^d} |\rho_\varepsilon(y)| \, \mathrm{d}y \leq M$ for some finite M;
- $\forall \, \delta > 0$, $\int_{\{|y| > \delta\}} |\rho_\varepsilon(y)| \, \mathrm{d}y \underset{\varepsilon \to 0}{\to} 0$.

The second condition is trivially satisfied in this case since our kernel is nonnegative. The last also follows easily since ρ_ε has compact support; more generally, this imposes rapid decay conditions on the tails of good kernels.

Since $\rho_\varepsilon \in C_c^\infty(\mathbb{R}^d)$, for "reasonable" f, $f_\varepsilon = \rho_\varepsilon * f \in C^\infty(\mathbb{R}^d)$ by the previous proposition. In fact, we'll see that in many contexts $f_\varepsilon \to f$ as $\varepsilon \to 0$ in some notion of convergence. So, f_ε provides a good, now smooth, approximation to f.

↪**Proposition 2.15**: Suppose $f \in L^\infty(\mathbb{R}^d)$ and f_ε is well-defined. Then, if f is continuous at x, then $f_\varepsilon(x) \to f(x)$ as $\varepsilon \to 0$.

If $f \in C(\mathbb{R}^d)$, then $f_\varepsilon \to f$ uniformly on compact sets.

Proof. f continuous at x gives that for every $\eta > 0$ there exists a $\delta > 0$ such that $|f(y) - f(x)| < \eta$ whenver $|x - y| < \delta$. Then

$$
\begin{aligned}
|f_\varepsilon(x) - f(x)| &= \left| \int \rho_\varepsilon(y) f(x-y) \,\mathrm{d}y - f(x) \underbrace{\int \rho_\varepsilon(y) \,\mathrm{d}y}_{=1} \right| \\
&= \left| \int \rho_\varepsilon(y) (f(x-y) - f(x)) \,\mathrm{d}y \right| \\
&\leq \int_{\{|y| \leq \delta\}} |f(x-y) - f(x)| |\rho_\varepsilon(y)| \,\mathrm{d}y + \int_{\{|y| > \delta\}} |f(x-y) - f(x)| |\rho_\varepsilon(y)| \,\mathrm{d}y \\
\left(\begin{smallmatrix}\text{cnty in first argument} \\ L^\infty\text{-bound in second}\end{smallmatrix}\right) \quad &\leq \int_{\{|y| \leq \delta\}} \eta |\rho_\varepsilon(y)| \,\mathrm{d}y + 2\|f\|_\infty \int_{|y| > \delta} |\rho_\varepsilon(y)| \,\mathrm{d}y \\
&\leq \eta \cdot M + 2\|f\|_\infty \int_{\{|y| > \delta\}} |\rho_\varepsilon|
\end{aligned}
$$

for $\varepsilon \to 0$, by using the second property of good kernels for the first bound. By the last property, the right-most term $\to 0$ as $\varepsilon \to 0$; moreover, then,

$$
\lim_{\varepsilon \to 0} |f_\varepsilon(x) - f(x)| \leq C\eta
$$

for some C and every $\eta > 0$, and thus $f_\varepsilon(x) \to f(x)$ as $\varepsilon \to 0$.

Now, if $f \in C(\mathbb{R}^d)$ fix a subset $K \subseteq \mathbb{R}^d$ compact. Hence, $\|f\|_{L^\infty(K)} < \infty$ and f uniformly continuous on K since K compact; so the modulus of continuity is uniform for all $x \in K$, so for $\delta > 0$ and for every $x \in K$,

$$
\int_{\{|y| \leq \delta\}} |f(x-y) - f(x)| |\rho_\varepsilon(y)| \,\mathrm{d}y \leq C\eta.
$$

Also, using the bound on f, we may write the second integral in the argument above as

$$
\int_{\varepsilon > |y| > \delta} |f(x-y) - f(x)| |\rho_\varepsilon(y)| \,\mathrm{d}y \leq \|f\|_{L^\infty(K + B_\varepsilon)} \int_{\{|y| > \delta\}} |\rho_\varepsilon(y)| \,\mathrm{d}y \underset{\varepsilon \to 0}{\to} 0
$$

where we take K slighly larger as $K + B_\varepsilon$, which is still compact. So, since this held for all $x \in K$,

$$
\max_{x \in K} |f_\varepsilon(x) - f(x)| \underset{\varepsilon \to 0}{\to} 0.
$$

Note that we proved the first for general good kernels but the second only in our constructed one. ■

Remark 2.27: This pointwise convergence result is why "good kernels" are called "approximations to the identity".

Remark 2.28: If $f \in C_c(\mathbb{R}^d)$, then $\operatorname{supp}(f_\varepsilon) \subseteq \overline{\operatorname{supp}\,(f) + B(0,\varepsilon)}$; so, f_ε is compactly supported if f is. Hence in this case $f_\varepsilon \to f$ uniformly on $\mathbb{R}^d$. More generally, there are many different restrictions one can place on the last claim, such as compact support of f, uniform continuity of f, compact support of the kernel, lack of compact support for the kernel but an L^∞ bound on f, etc. In practice, the proofs are all the same, with different bounds; namely one finds something of the form

$$|f_\varepsilon(x) - f(x)| \leq \underbrace{\int_{|y|\,<\delta} (\dots)}_{\substack{\text{small by} \\ \text{(uniform) continuity}}} + \underbrace{\int_{|y|\,\geq\delta} (\dots)}_{\substack{\text{small by} \\ \text{compact support, etc}}}$$

↪**Theorem 2.26** (Weierstrass Approximation Theorem): Let $[a,b] \subseteq \mathbb{R}$ and let $f \in C([a,b])$. Then for every $\eta > 0$, there exists a polynomial $P_N(x)$ of degree N such that

$$\|P_N - f\|_{L^\infty([a,b])} < \eta.$$

That is, polynomials are dense in $C([a,b])$.

Proof. Extend f to be continuous with compact support on all of $\mathbb{R}$ in whatever convenient way, such that $\operatorname{supp}(f) \subseteq [-M, M]$ for some sufficiently large $M > 0$. Consider now

$$K_\varepsilon(x) := \frac{1}{\sqrt{\varepsilon}} e^{-\frac{\pi x^2}{\varepsilon}},$$

noting that

$$\int_{-\infty}^{\infty} K_\varepsilon(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} \frac{1}{\sqrt{\varepsilon}} e^{-\frac{\pi x^2}{\varepsilon}}\,\mathrm{d}x = 1,$$

which is clear by a change of variables $y = \frac{\sqrt{2\pi}}{\sqrt{\varepsilon}} x$. As a consequence, $\int_{-\infty}^{\infty} |K_\varepsilon(x)|\,\mathrm{d}x = 1 < \infty$, since $K_\varepsilon \geq 0$. Finally,

$$\begin{aligned}
\int_{|x|>\delta} |K_\varepsilon(x)|\,\mathrm{d}x &= \int_{|x|>\delta} \frac{1}{\sqrt{\pi}} e^{-\frac{\pi x^2}{\varepsilon}}\,\mathrm{d}x \\
&= \int_{|y|>\frac{\sqrt{2\pi}}{\sqrt{\varepsilon}}\delta} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}y \\
\text{since } |y| \geq 1 \text{ here for suff. small } \varepsilon \qquad &\leq \int_{|y|>\frac{\sqrt{2\pi}}{\sqrt{\varepsilon}}\delta} \frac{|y|}{\sqrt{2\pi}} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}y \\
&\leq C e^{-\frac{y^2}{2}} \Big|_{\frac{\sqrt{2\pi}}{\sqrt{\varepsilon}}\delta}^{\infty} \underset{\varepsilon\to 0}{\to} 0.
\end{aligned}$$

So, K_ε is a good kernel, and so $(f * K_\varepsilon)(\varepsilon) \underset{\varepsilon\to 0}{\to} f$ uniformly in $[a,b]$ by our last remark. In particular, for $\eta > 0$ there is some $\varepsilon_0 > 0$,

$$\left\|\left(f * K_{\varepsilon_0}\right) - f\right\|_{L^\infty([a,b])} < \frac{\eta}{2}.$$

We claim now that there is a polynomial P_N such that $\left\|P_N - \left(f * K_{\varepsilon_0}\right)\right\|_{L^\infty([a,b])} < \frac{\eta}{2}$. Recall that $e^x = \sum_{n=0}^\infty \frac{x^n}{n!}$, which converges uniformly on compact sets. So, there exists a polynomial S_N (from truncating this sum) such that $\left\|K_{\varepsilon_0} - S_N\right\|_{L^\infty([-M,M])} < \frac{\eta}{4\|f\|_\infty M}$. Thus,

$$\begin{aligned} \left|f * K_{\varepsilon_0}(x) - f * S_n(x)\right| &\leq \left|\int f(x-y)\left(K_{\varepsilon_0}(y) - S_N(y)\right) \mathrm{d}y\right| \\ \operatorname{supp}(f) \subset [-M,M] \qquad &\leq \int_{-M}^{M} |f(x-y)|\left|K_{\varepsilon_0}(y) - S_N(y)\right| \mathrm{d}y \\ &\leq 2M\|f\|_\infty \frac{\eta}{4M\|f\|_\infty} = \frac{\eta}{2}, \end{aligned}$$

for every x. Let $P_N(x) = (f * S_n)(x)$, which we see to be a polynomial. ■

↪**Theorem 2.27**: Let $f \in L^p(\mathbb{R}^d)$ with $p \in [1,\infty)$. Then $f_\varepsilon \underset{L^p(\mathbb{R}^d)}{\to} f$.

Proof. Since $f \in L^p(\mathbb{R}^d)$, for every $\eta > 0$ there is a $\tilde{f} \in C_c(\mathbb{R}^d)$ such that $\left\|f - \tilde{f}\right\|_p < \eta$. Since $\tilde{f} \in C_c(\mathbb{R}^d)$, by the previous theorem dealing with mollifiers and uniform convergence, $\tilde{f}_\varepsilon \to \tilde{f}$ uniformly. In particular, we have $\left\|\tilde{f}_\varepsilon - \tilde{f}\right\|_p \underset{p}{\to} 0$, hence

$$\|f - f_\varepsilon\|_p \leq \left\|f_\varepsilon - \tilde{f}_\varepsilon\right\|_p + \left\|\tilde{f}_\varepsilon - \tilde{f}\right\|_p + \left\|\tilde{f} - f\right\|_p.$$

We've dealt with the second two bounds. For the first,

$$\begin{aligned} \left\|f_\varepsilon - \tilde{f}_\varepsilon\right\|_p &= \left\|\left(f - \tilde{f}\right) * \rho_\varepsilon\right\|_p \\ \text{(Young's)} \qquad &\leq \|\rho_\varepsilon\|_1 \left\|f - \tilde{f}\right\|_p = \left\|f - \tilde{f}\right\|_p, \end{aligned}$$

so

$$\|f - f_\varepsilon\|_p \leq 2\left\|f - \tilde{f}\right\|_p + \left\|\tilde{f}_\varepsilon - \tilde{f}\right\|_p < 3\eta.$$

■

↪**Corollary 2.6**: $C_c^\infty(\mathbb{R}^d)$ dense in $L^p(\mathbb{R}^d)$.

Proof. We showed $\tilde{f}_\varepsilon$ approximates f in $L^p(\mathbb{R}^d)$, and by construction $\tilde{f}_\varepsilon$ is smooth with compact support. ■

§2.12 Strong Compactness in $L^p(\mathbb{R}^d)$

We saw that for $p \in (1,\infty)$, $\{f_n\} \subset L^p(\Omega)$, that any bounded sequence admits a weakly converging subsequence, $f_{n_k} \underset{L^p}{\rightharpoonup} f$. In addition, if the norms also converge i.e. $\lim_{n\to\infty} \|f_n\|_p = \|f\|_p$, then we actually have strong convergence $f_{n_k} \underset{L^p}{\to} f$.

We provide now a strong compactness result in L^p, akin to Arzelà-Ascoli.

↪**Theorem 2.28** (Strong Compactness): Let $\{f_n\} \subseteq L^p(\mathbb{R}^d)$ for $p \in [1,\infty)$ s.t.
i. $\exists C > 0$ s.t. $\|f_n\|_p < C \,\forall\, n$, i.e. $\{f_n\}$ uniformly bounded in L^p;
ii. $\lim_{|h|\to 0} \|f_n - \tau_h f_n\|_p = 0$ uniformly in n, i.e. for every $\eta > 0$, there exists $\delta > 0$ such that if $|h| < \delta$, $\int |f_n(x) - f_n(x-h)|^p \,\mathrm{d}x < \eta^p$ *for every* n;

Then, for any $\Omega \subseteq \mathbb{R}^d$ with finite measure, there exists a subsequence $\{f_{n_k}\}$ such that $f_{n_k} \underset{L^p(\Omega)}{\to} f$.

Proof. Recall that $L^p(\Omega)$ is a complete metric space, so TFAE:
1. sequential compactness;
2. totally bounded (& complete);
3. compact.

Let $\mathcal{F} = \{f \in L^p(\mathbb{R}^d)$ satisfying i., ii.$\}$ and fix $\Omega \subseteq \mathbb{R}^d$ with finite measure. We aim to show that $\mathcal{F}|_\Omega$ is sequentially compact in $L^p(\Omega)$ (with no regard to whether the limit lives in $\mathcal{F}_\Omega$); equivalently, we wish to show $\mathcal{F}|_\Omega$ is precompact in $L^p(\Omega)$ i.e. $\overline{\mathcal{F}|_\Omega}$ is compact. Since $\overline{\mathcal{F}|_\Omega}$ is a complete metric space, to prove the result it suffices to show that $\mathcal{F}|_\Omega$ is totally bounded (recall: for every $\delta > 0$, $\mathcal{F}|_\Omega \subseteq \bigcup_{i=1}^N B_{L^p(\Omega)}(g_i, \delta)$). We'll do this using mollifiers and AA.

Step 1: Fix η, δ as in ii. in the statement of the theorem, and let $f \in \mathcal{F}$. Then, for every $\varepsilon < \delta$, we claim

$$\|(\rho_\varepsilon * f) - f\|_{L^p(\mathbb{R}^d)} < \eta.$$

We have

$$\begin{aligned}
|(\rho_\varepsilon * f)(x) - f(x)| &= \left| \int_{B_\varepsilon} \rho_\varepsilon(y) f(x-y)\,\mathrm{d}y - f(x) \int \rho_\varepsilon(y)\,\mathrm{d}y \right| \\
&\le \int_{B_\varepsilon} \rho_\varepsilon(y) |f(x-y) - f(x)|\,\mathrm{d}y \\
&= \int_{B_\varepsilon} \rho_\varepsilon^{\frac{1}{q}}(y) \rho_\varepsilon^{\frac{1}{p}}(y) |f(x-y) - f(x)|\,\mathrm{d}y \\
\text{(Holder's)} \qquad &\le \left(\int \rho_\varepsilon(y) |f(x-y) - f(x)|^p \,\mathrm{d}y \right)^{1/p} \underbrace{\left(\int \rho_\varepsilon(y)\,\mathrm{d}y \right)^{1/q}}_{=1},
\end{aligned}$$

and hence

$$
\begin{aligned}
\int |(\rho_\varepsilon * f)(x) - f(x)|^p \,\mathrm{d}x &\leq \iint \rho_\varepsilon(y) |f(x-y) - f(x)|^p \,\mathrm{d}y \,\mathrm{d}x \\
\text{(Tonelli's)} \qquad &= \int_{B_\varepsilon} \rho_\varepsilon(y) \underbrace{\int |f(x-y) - f(x)|^p \,\mathrm{d}x \,\mathrm{d}y}_{\varepsilon < \delta \Rightarrow \eta^p} \\
&< \eta^p \underbrace{\int_{B_\varepsilon} \rho_\varepsilon(y) \,\mathrm{d}y}_{=1} = \eta^p,
\end{aligned}
$$

hence $\|(\rho_\varepsilon * f)(x) - f(x)\|_p < \eta$.

Step 2: We first claim that there exists some C_ε such that for any $f \in \mathcal{F}$,

$$
\|\rho_\varepsilon * f\|_\infty \leq C_\varepsilon \|f\|_p, \qquad (1)
$$

and that for any $x_1, x_2 \in \mathbb{R}^d$,

$$
|(\rho_\varepsilon * f)(x_1) - (\rho_\varepsilon * f)(x_2)| \leq C_\varepsilon \|f\|_p |x_1 - x_2|. \qquad (2)
$$

In particular, this shows that *for ε fixed*, $(\rho_\varepsilon * f)$ satisfy hypothesis of AA. Remark that the first is a uniform boundedness type condition for $\rho_\varepsilon * f$, and the second is a uniform Lipschitz bound.

For the first claim (1),

$$
\begin{aligned}
|(\rho_\varepsilon * f)(x)| &= \left| \int \rho_\varepsilon(x-y) f(y) \,\mathrm{d}y \right| \\
\text{(Holder's)} \qquad &\leq \left(\int \rho_\varepsilon^q(x-y) \,\mathrm{d}y \right)^{\frac{1}{q}} \cdot \|f\|_p \\
&= \|\rho_\varepsilon\|_q \|f\|_p,
\end{aligned}
$$

so we have the bound with $C_\varepsilon := \|\rho_\varepsilon\|_q$ since the bound is independent of x.

Remark 2.29: One can explicitly compute $\|\rho_\varepsilon\|_q$, and realize that it will in general depend explicitly on ε.

For the second statement (2), we find that $\nabla(\rho_\varepsilon * f) = (\nabla \rho_\varepsilon) * f$ since the RHS is finite, because

$$
(\nabla \rho_\varepsilon * f)(x) = \int \nabla \rho_\varepsilon(x-y) f(y) \,\mathrm{d}y \leq \|\nabla \rho_\varepsilon\|_q \|f\|_p.
$$

So,

$$
\|\nabla(\rho_\varepsilon * f)\|_\infty \leq \underbrace{\|\nabla \rho_\varepsilon\|_q}_{=:C_\varepsilon} \|f\|_p.
$$

By the mean-value theorem then, we have all together

$$
\begin{aligned}
\|(\rho_\varepsilon * f)(x_1) - (\rho_\varepsilon * f)(x_2)\| &\leq \|\nabla(\rho_\varepsilon * f)\|_\infty |x_1 - x_2| \\
&\leq C_\varepsilon \|f\|_p |x_1 - x_2|.
\end{aligned}
$$

This proves (2).

Step 3: Next, we claim that for $\eta > 0$ and fixed $\varepsilon < \eta$ and $\Omega \subseteq \mathbb{R}^d$ with finite measure, there exists $E \subseteq \Omega \subseteq \mathbb{R}^d$ such that E is bounded, i.e. $E \subseteq B(0, M)$ where M sufficiently large, and moreover that $\|f\|_{L^p(\Omega \setminus E)} < \eta$ for every $f \in \mathcal{F}$.

We have that

$$
\|f\|_{L^p(\Omega \setminus E)} \leq \|f - (\rho_\varepsilon * f)\|_{L^p(\mathbb{R}^d)} + \|\rho_\varepsilon * f\|_{L^p(\Omega \setminus E)}.
$$

By the very first step of the proof, the first term is $< \eta$, so this is bounded by

$$
\begin{aligned}
&< \eta + \left(\int_{\Omega / E} |\rho_\varepsilon * f|^p \, \mathrm{d}x \right)^{1/p} \\
&< \eta + \|\rho_\varepsilon * f\|_\infty \; |\Omega \setminus E|^{\frac{1}{p}} \\
&< \eta + C_\varepsilon \|f\|_p |\Omega \setminus E|^{\frac{1}{p}}.
\end{aligned}
$$

C_ε finite and $\|f\|_p$ upper bounded uniformly over $\mathcal{F}$, so it suffices to construct E with the measure of $\Omega \setminus E$ sufficiently small, so we can get $\|f\|_{L^p(\Omega \setminus E)} < 2\eta$.

Step 4: Fix $\eta > 0$. We claim $\mathcal{F}|_\Omega$ is totally bounded. Let $\varepsilon < \delta$ then let

$$
\mathcal{H} := (\rho_\varepsilon * \mathcal{F})|_{\overline{E}} = \{\rho_\varepsilon * f|_E : f \in \mathcal{F}\}.
$$

$E \subseteq \Omega \subseteq \mathbb{R}^d$ is bounded implies $\overline{E}$ is compact. So by Step 2., we showed $(\rho_\varepsilon * \mathcal{F})$ satisfies hypotheses of AA on $\overline{E}$. Hence, $\mathcal{H}$ is precompact in $C(\overline{E})$. Thus, since we have uniform convergence we certainly have L^p convergence thus $\mathcal{H}$ also precompact in $L^p(\overline{E})$. Thus, for $\eta > 0$, there exists $\{\overline{g}_i\} \subseteq L^p(\overline{E})$ such that

$$
\mathcal{H} \subseteq \bigcup_{i=1}^N B_{L^p(\overline{E})}(\overline{g}_i, \eta). \qquad \star
$$

Let $g_i : \Omega \to \mathbb{R}$ be given by

$$
g_i(x) = \begin{cases} \overline{g}_i & \text{on } E \\ 0 & \text{on } \Omega \setminus E \end{cases}.
$$

Then, we claim $\mathcal{F}|_\Omega \subseteq \bigcup_{i=1}^N B_{L^p(\Omega)}(g_i, 3\eta)$. If $f \in \mathcal{F}$ by $\star$, there is an i such that $\left\|\rho_\varepsilon * f - \overline{g}_i\right\|_{L^p(\overline{E})} < \eta$. But also,

$$
\begin{aligned}
\|f - g_i\|_{L^p(\Omega)}^p &= \int_{\Omega\setminus E} |f|^p \,\mathrm{d}x + \int_{\overline{E}} \left|f - \overline{g}_i\right|^p \mathrm{d}x \\
&= \|f\|_{L^p(\Omega\setminus E)}^p + \int_{\overline{E}} \left|f - \overline{g}_i\right|^p \mathrm{d}x \\
\text{(Step 3.)} \qquad &\le \eta^p + \int_{\overline{E}} \left|f - \overline{g}_i\right|^p \mathrm{d}x.
\end{aligned}
$$

Recall $(a+b)^{\frac{1}{p}} \le a^{\frac{1}{p}} + b^{\frac{1}{p}}$. Applying this bound to the above, we find

$$
\begin{aligned}
\|f - g_i\|_{L^p(\Omega)} &\le \eta + \left\|f - \overline{g}_i\right\|_{L^p(\overline{E})} \\
&\le \eta + \underbrace{\|f - f * \rho_\varepsilon\|_{L^p(\mathbb{R}^d)}}_{<\eta \text{ by Step 1.}} + \underbrace{\left\|(f * \rho_\varepsilon) - \overline{g}_i\right\|_{L^p(\overline{E})}}_{<\eta \text{ by } \star} \\
&\le 3\eta.
\end{aligned}
$$

Hence, $\mathcal{F}|_\Omega \subseteq \bigcup_{i=1}^N B(g_i, 3\eta)$, thus $\mathcal{F}|_\Omega$ is sequentially compact so any sequence in $\mathcal{F}$ has a converging subsequence, which proves the theorem. ■

Remark 2.30: This can be extended to $L^p(\mathbb{R}^d)$ with some conditions.

§3 Introduction to Fourier Analysis

References are Folland, Chapter 8 and *Fourier Analysis* by Stein & Sharkarchi.

§3.1 Fourier Series

We will denote the torus $\mathbb{T} = [0,1) \simeq \mathbb{R}/\mathbb{Z}$ (with 1 identified back with 0), and specifically complex-valued functions on the torus

$$
L^2(\mathbb{T}) = \left\{ f : \mathbb{T} \to \mathbb{C} \;\middle|\; \int_0^1 |f(x)|^2 \,\mathrm{d}x < \infty \right\},
$$

where now $|\cdot|$ the modulus (i.e. $|a+bi|^2 = a^2 + b^2$). Equivalently, $f : \mathbb{T} \to \mathbb{C}$ can be identified with $\tilde{f} : \mathbb{R} \to \mathbb{C}$ which is periodic.

↪**Proposition 3.1**: The function $L^2(\mathbb{T}) \times L^2(\mathbb{T}) \to \mathbb{C}$

$$
(f, g) = \int_0^1 f(x)\overline{g(x)} \,\mathrm{d}x
$$

is an inner product on $L^2(\mathbb{T})$. In particular, $(L^2(\mathbb{T}), (\cdot,\cdot))$ a Hilbert space.

Proof. For $\mathbb{C}$-valued functions, we need to check:

- linearity in the first variable: for $\alpha \in \mathbb{C}$,

$$
(\alpha f + h, g) = \int_0^1 (\alpha f + h)\overline{g} \,\mathrm{d}x = \alpha(f, g) + (h, g)
$$

by linearity of the integral;

- conjugate symmetry:

$$\begin{aligned}
\int_0^1 f(x)\overline{g(x)}\,\mathrm{d}x &= \int_0^1 (\mathrm{Re}(f)+i\mathrm{Im}(f))(\mathrm{Re}(g)-i\mathrm{Im}(g))\,\mathrm{d}x \\
&=: \int_0^1 (a+ib)(c-id)\,\mathrm{d}x \\
&= \int_0^1 (ac+bd)+i(bc-ad)\,\mathrm{d}x \\
&= \int_0^1 (ac+bd)-i(ad-bc)\,\mathrm{d}x \\
&= \overline{\int_0^1 g\overline{f}\,\mathrm{d}x} = \overline{(g,f)};
\end{aligned}$$

- f inner product with f properties:

$$(f,f) = \int_0^1 f(x)\overline{f(x)}\,\mathrm{d}x = \int_0^1 |f(x)|^2\,\mathrm{d}x = \|f\|^2_{L^2(\mathbb{T})} \geq 0, = 0 \text{ iff } f \equiv 0.$$

We know $L^2(\mathbb{T})$ is complete, so $L^2(\mathbb{T})$ a Hilbert space with this inner product since it induces the same norm as the usual norm L^2-norm. ■

↪**Theorem 3.1**: Let $e_n(x) := e^{2\pi i n x}$ for $n \in \mathbb{Z}$. Then, $\{e_n\}_{n\in\mathbb{Z}}$ is an orthonormal basis of $L^2(\mathbb{T})$.

PROOF. For orthonormality, if $n \neq m$,

$$\begin{aligned}
(e_n, e_m) &= \int_0^1 e^{2\pi i n x} e^{-2\pi i m x}\,\mathrm{d}x \\
&= \int_0^1 e^{2\pi i (n-m)x}\,\mathrm{d}x \\
&= \frac{1}{2\pi i(n-m)} e^{2\pi i(n-m)x}\Big|_0^1 \\
&= \frac{1}{2\pi i(n-m)}\left[e^{2\pi i(n-m)} - 1\right] \\
&= \frac{1}{2\pi i(n-m)}\left[\underbrace{\cos(2\pi(n-m))}_{=1} + \underbrace{i\sin(2\pi(n-m))}_{=0} - 1\right] = 0,
\end{aligned}$$

and if $n = m$,

$$(e_n, e_n) = \int_0^1 \left|e^{2\pi i n x}\right|^2\,\mathrm{d}x = \int_0^1 1\,\mathrm{d}x = 1.$$

To prove its a basis, we use Stone-Weierstrass. $\mathbb{T}$ is compact; let

$$\mathcal{A} := \left\{ \sum_{n=-N}^{N} \alpha_n e_n : \alpha_n \in \mathbb{C}, N \in \mathbb{N} \right\}.$$

Notice $e_n e_m = e^{2\pi i(n+m)x} = e_{n+m}$, and $e_0 = 1$, so this family stays closed under multiplication (and clearly addition and scalar multiplication by definition), so is an algebra which contains constant functions. Also, if $x_1 \neq x_2$ and $x_1, x_2 \in [0, 1)$, then if $n \neq 0$, $e_n(x_1) = e^{2\pi i n x_1} \neq e^{2\pi i n x_2} = e_n(x_2)$, so $\mathcal{A}$ separates points. By (complex) Stone-Weierstrass, then we know $\mathcal{A}$ is dense in $C(\mathbb{T}, \mathbb{C})$ with respect to $\|\cdot\|_\infty$. We know $C(\mathbb{T}, \mathbb{C})$ is dense in $L^2(\mathbb{T})$ (by some mollifier argument, for example) with respect to $\|\cdot\|_{L^2(\mathbb{T})}$. So,

$$f(x) = \lim_{N\to\infty} \sum_{n=-N}^{N} \alpha_n e_n(x),$$

with the limit taken in the sense of $L^2(\mathbb{T})$. ■

Recall that in Hilbert spaces, TFAE:

- $\{e_n\}$ are a basis, i.e. $f = \sum_{n=-\infty}^{\infty} \alpha_n e_n = \sum_{n=-\infty}^{\infty} (f, e_n) e_n$, in $L^2(\mathbb{T})$;
- if $(f, e_n) = 0$ for every n, $f \equiv 0$ (completeness);
- $\|f\|_{L^2(\mathbb{T})}^2 = \sum_{n=-\infty}^{\infty} |(f, e_n)|^2$ (Parseval's).

With this in mind, we define:

↪**Definition 3.1** (Fourier Series): Let

$$\hat{f}(n) := (f, e_n) = \int_0^1 f(x) e^{-2\pi i n x} \, \mathrm{d}x.$$

Then, the *complex Fourier series* is defined by

$$\sum_{n=-\infty}^{\infty} \hat{f}(n) e^{2\pi i n x}.$$

Remark 3.1: A Fourier series can be defined for any periodic function, while we only do so for 1-periodic here. If f were L-periodic, we'd define

$$\hat{f}_L(n) := \frac{1}{L} \int_0^L f(x) e^{\frac{-2\pi i n x}{L}} \, \mathrm{d}x,$$

with complex Fourier series $\sum_{n=-\infty}^{\infty} \hat{f}_L(n) e^{\frac{2\pi i n x}{L}}$.

Remark 3.2: We can also make Fourier series to be real-valued, with sines and cosines, of the form

$$A_0 + \sum_{n=1}^{\infty}\left[A_n \cos\left(\frac{2n\pi x}{L}\right) + B_n \sin\left(\frac{2n\pi x}{L}\right)\right],$$

for some A_n, B_n also given by inner products.

What conditions do we need on f to make this series converge? In the general L^2-theory, we just need $f \in L^2(\mathbb{T})$. By Parseval's,

$$\|f\|^2_{L^2(\mathbb{T})} = \sum_{n=-\infty}^{\infty} \left|\hat{f}(n)\right|^2.$$

So, the operator $\hat{\cdot} : L^2(\mathbb{T}) \to \ell^2(\mathbb{C})$. Note that this implies $\lim_{n\to\infty} \left|\hat{f}(n)\right|^2 = 0$, so also $\lim_{n\to\infty}\left|\hat{f}(n)\right| = 0$. This proves the following proposition:

↪**Proposition 3.2** (Riemann-Lebesgue Lemma): If $f \in L^2(\mathbb{T})$,

$$\lim_{n\to\infty}\left|\hat{f}(n)\right| = 0.$$

Remark 3.3: This result in *very* useful, particularly for the real Fourier Series. In particular, it tells us statements such as

$$\lim_{n\to\infty} \int_0^1 f(x)\sin(2n\pi x)\,\mathrm{d}x = 0,$$

with similar for the cosine term. These are so-called "oscillatory integrals".

While the $L^2(\mathbb{T})$-theory is very useful for Hilbert space interpretation, we are really concerned with the partial sums

$$S_N(x) = \sum_{n=-N}^{N} \hat{f}(n)e^{2\pi i n x},$$

and ways it might converge. We may rewrite by definition

$$\begin{aligned} S_N(x) &= \sum_{n=-N}^{N}\left(\int_0^1 f(y)e^{-2\pi i n y}\,\mathrm{d}y\right)e^{2\pi i n x} \\ \text{(because finite sum)} \qquad &= \int_0^1 f(y)\sum_{n=-N}^{N} e^{2\pi i n(x-y)}\,\mathrm{d}y \\ (* \text{ just over } [0,1)) \qquad &= (f * D_N)(x), \qquad D_N(x) := \sum_{n=-N}^{N} e^{2\pi i n x}. \end{aligned}$$

So in short,

$$S_N(x) = (f * D_N)(x),$$

where $D_N(x)$ is called the *Dirichlet kernel*. Let's look at some of its properties.

$$D_N(x) = 1 + \sum_{n=1}^{N} [e^{2\pi i n x} + e^{-2\pi i n x}],$$

so

$$\int_0^1 D_N(x)\, \mathrm{d}x = \int_0^1 1\, \mathrm{d}x + \underbrace{\sum_{n=1}^{N} \int_0^1 (\text{some periodic functions})}_{=0} = 1,$$

by periodicity. However, $D_N(x)$ is not actually a good kernel; one can show that $\int_0^1 |D_N(x)|\, \mathrm{d}x \geq C \log N$ as $N \to \infty$.

Note too that

$$\begin{aligned}
D_N(x) &= \sum_{n=-N}^{N} e^{2\pi i n x} \\
&= \sum_{n=0}^{2N} e^{2\pi i (n-N) x} \\
&= e^{-2\pi i N x} \sum_{n=0}^{2N} \left(e^{2\pi i x}\right)^n \\
&= e^{-2\pi i N x} \left(\frac{1 - e^{2\pi i (2N+1)} x}{1 - e^{2\pi i x}} \right) \qquad\qquad \text{(geometric series)} \\
&= \frac{e^{-2\pi i N x} - e^{2\pi i (N+1) x}}{1 - e^{2\pi i x}} \cdot \frac{e^{-2\pi i \frac{x}{2}}}{e^{-2\pi i \frac{x}{2}}} \\
&= \frac{e^{-2\pi i (N + \frac{1}{2}) x} - e^{2\pi i (N + \frac{1}{2}) x}}{e^{-2\pi i \frac{x}{2}} - e^{2\pi i \frac{x}{2}}} \\
&= \frac{\sin\left(2\pi \left(N + \frac{1}{2}\right) x\right)}{\sin\left(2\pi \frac{x}{2}\right)}.
\end{aligned}$$

This form leads nicely to the following results.

↪**Theorem 3.2** (Pointwise Convergence): Let $f \in L^2(\mathbb{T})$ and suppose f is Lipschitz continuous at x_0. Then,

$$S_N(x_0) \to f(x_0).$$

Proof. Left as an exercise.

■

↪**Theorem 3.3** (Uniform convergence): If $f \in C^2(\mathbb{T})$, then $S_N(x) \to f(x)$ uniformly on $\mathbb{T}$.

Proof. Exercise. ■

Remark 3.4: In fact, we see that $\hat{f}(n) = \int_0^1 f(x)e^{-2\pi i n x}\,\mathrm{d}x$ is well-defined whenever $f \in L^1(\mathbb{T})$. So, we can view

$$\hat{\cdot} : L^1(\mathbb{T}) \to \ell^\infty(\mathbb{C}).$$

Remark 3.5: All the prior results can be extended to $f \in L^1(\mathbb{T})$, via density.

§3.2 Introduction to the Fourier Transform

↪**Definition 3.2** (Fourier Transform): Let $f : \mathbb{R} \to \mathbb{C}$. Then, for any $\zeta \in \mathbb{R}$, define

$$\hat{f}(\zeta) := \int_{\mathbb{R}} f(x)e^{-2\pi i \zeta x}\,\mathrm{d}x.$$

Remark 3.6: If $f \in L^1(\mathbb{R})$,

$$\left|\hat{f}(\zeta)\right| \leq \int_{-\infty}^{\infty} |f(x)| \underbrace{\left|e^{-2\pi i \zeta x}\right|}_{=1}\,\mathrm{d}x = \|f\|_{L^1(\mathbb{R})}$$

so in particular, $\hat{f} \in L^\infty(\mathbb{R})$. Moreover,

$$\begin{aligned} \left|\hat{f}(\zeta + h) - \hat{f}(\zeta)\right| &= \left|\int_{\mathbb{R}} f(x)e^{-2\pi i(\zeta+h)x} - f(x)e^{-2\pi i \zeta x}\,\mathrm{d}x\right| \\ &= \left|\int_{\mathbb{R}} f(x)e^{-2\pi i \zeta x}\left(e^{-2\pi i h x} - 1\right)\right| \\ &\leq \int_{\mathbb{R}} |f(x)|\left|e^{-2\pi i h x} - 1\right|\,\mathrm{d}x. \end{aligned}$$

We have that

$$\lim_{h\to 0}\left|e^{-2\pi i h x} - 1\right| = 0$$

for a.e. $x \in \mathbb{R}$, and

$$\int_{\mathbb{R}} |f(x)|\left|e^{-2\pi i h x} - 1\right|\,\mathrm{d}x \leq 2\int_{\mathbb{R}} |f(x)|\,\mathrm{d}x = 2\|f\|_{L^1(\mathbb{R})},$$

so we can apply dominated convergence theorem to find

$$\lim_{h\to 0}\left|\hat{f}(\zeta + h) - \hat{f}(\zeta)\right| \leq \int_{\mathbb{R}} |f(x)| \underbrace{\lim_{h\to 0}\left|e^{-2\pi i h x} - 1\right|}_{=0}\,\mathrm{d}x = 0,$$

so $\hat{f} \in C(\mathbb{R})$.

↪**Proposition 3.3** (Properties of the Fourier Transform): Let $f \in L^1(\mathbb{R})$ and $g \in L^1(\mathbb{R})$. Then,
a. $\widehat{\tau_y f}(\zeta) = e^{-2\pi i \zeta y}\hat{f}(\zeta)$, and $\tau_\eta \hat{f}(\zeta) = \widehat{e^{2\pi i \eta (\cdot)} f(\cdot)}(\zeta)$;
b. $\widehat{f * g} = \hat{f} \cdot \hat{g}$;
c. $\int_{\mathbb{R}} f(x)\hat{g}(x)\,\mathrm{d}x = \int_{\mathbb{R}} \hat{f}(x) g(x)\,\mathrm{d}x$.

Proof. a. A change of variables gives

$$\widehat{\tau_y f}(\zeta) = \int_{\mathbb{R}} f(x-y)e^{-2\pi i \zeta x}\,\mathrm{d}x = e^{-2\pi i \zeta y}\int_{\mathbb{R}} f(x)e^{-2\pi i \zeta x}\,\mathrm{d}x = e^{-2\pi i \zeta y}\hat{f}(\zeta).$$

Similarly,

$$\tau_\eta \hat{f}(\zeta) = \int_{\mathbb{R}} f(x)e^{-2\pi i(\zeta-\eta)x}\,\mathrm{d}x = \int_{\mathbb{R}} f(x)e^{2\pi i \eta x}\cdot e^{-2\pi i \zeta x}\,\mathrm{d}x = \widehat{e^{2\pi i \eta(\cdot)} f(\cdot)}(\zeta).$$

b. First, by Young's inequality $f * g \in L^1(\mathbb{R})$ so this makes sense. Moreover, since $f, g \in L^1(\mathbb{R})$, everything we need to be finite is finite, so we can apply Fubini's theorem to find

$$\begin{aligned}
\widehat{f * g}(\zeta) &= \int\left(\int f(x-y)g(y)\,\mathrm{d}y\right)e^{-2\pi i \zeta x}\,\mathrm{d}x \\
&= \int\left(\int f(x-y)e^{-2\pi i \zeta x}\,\mathrm{d}x\right)g(y)\,\mathrm{d}y \\
&= \int\left(\int f(x-y)e^{-2\pi i \zeta (x-y)}\,\mathrm{d}x\right)e^{-2\pi \zeta y}g(y)\,\mathrm{d}y \\
&= \left(\int f(x)e^{-2\pi i \zeta x}\,\mathrm{d}x\right)\left(\int g(y)e^{-2\pi i \zeta y}\,\mathrm{d}y\right) = \hat{f}(\zeta)\cdot\hat{g}(\zeta),
\end{aligned}$$

where we "multiply by 1" in the second to last line to change variables in the appropriate way.

c. We can apply Fubini's again,

$$\begin{aligned}
\int f(x)\hat{g}(x)\,\mathrm{d}x &= \int f(x)\left(\int g(y)e^{-2\pi i x y}\,\mathrm{d}y\right)\mathrm{d}x \\
&= \int g(y)\left(\int f(x)e^{-2\pi i x y}\,\mathrm{d}x\right)\mathrm{d}y \\
&= \int g(y)\hat{f}(y)\,\mathrm{d}y.
\end{aligned}$$

■

↪**Lemma 3.1**: Let $f(x) = e^{-\pi a x^2}$ for $a > 0$. Then,

$$\hat{f}(\zeta) = \frac{1}{\sqrt{a}}e^{-\pi\frac{\zeta^2}{a}}.$$

PROOF. First, note that

$$\widehat{\frac{\mathrm{d}}{\mathrm{d}x}f}(\zeta) = \int_{-\infty}^{\infty} f'(x)e^{-2\pi i\zeta x}\,\mathrm{d}x$$

$$= f(x)e^{-2\pi i\zeta x}\Big|_{x=-\infty}^{\infty} - \int_{-\infty}^{\infty} f(x)(-2\pi i\zeta)e^{-2\pi i\zeta x}\,\mathrm{d}x.$$

Specifying $f(x) = e^{-2\pi a x^2}$, this becomes

$$= e^{-\pi a x^2}\cdot e^{-2\pi i\zeta x}\Big|_{x=-\infty}^{\infty} - \int_{-\infty}^{\infty} e^{-\pi a x^2}(-2\pi i\zeta)e^{-2\pi i\zeta x}\,\mathrm{d}x$$

$$= 2\pi i\zeta\cdot\int_{-\infty}^{\infty} e^{-\pi a x^2}e^{-2\pi i\zeta x}\,\mathrm{d}x = 2\pi i\zeta\cdot\hat{f}(\zeta).$$

On the other hand,

$$\frac{\mathrm{d}}{\mathrm{d}\zeta}\hat{f}(\zeta) = \frac{\mathrm{d}}{\mathrm{d}\zeta}\left(\int_{-\infty}^{\infty} f(x)e^{-2\pi i\zeta x}\,\mathrm{d}x\right)$$

$$= \int_{-\infty}^{\infty} f(x)(-2\pi i x)e^{-2\pi i\zeta x}\,\mathrm{d}x,$$

assuming finiteness; indeed,

$$\left|\int_{-\infty}^{\infty} e^{-\pi a x^2}(-2\pi x)e^{-2\pi i\zeta x}\,\mathrm{d}x\right| \le C\int_{-\infty}^{\infty}|x|e^{-\pi a x^2}\,\mathrm{d}x$$

$$= 2C\int_{0}^{\infty} x e^{-\pi a x^2}\,\mathrm{d}x = Ce^{-\pi a x^2}\Big|_{0}^{\infty} < \infty,$$

so our differentiation was valid. Thus, combining these two,

$$\frac{\mathrm{d}}{\mathrm{d}\zeta}\hat{f}(\zeta) = \int_{-\infty}^{\infty} -2\pi i x f(x)e^{-2\pi i\zeta x}\,\mathrm{d}x$$

$$= \int_{-\infty}^{\infty} i\left(-2\pi x e^{-\pi a x^2}\right)e^{-2\pi i\zeta x}\,\mathrm{d}x$$

$$= \int_{-\infty}^{\infty} \frac{i}{a}f'(x)e^{-2\pi i\zeta x}\,\mathrm{d}x$$

$$= \frac{i}{a}2\pi i\zeta\hat{f}(\zeta)$$

$$\Rightarrow \frac{\mathrm{d}}{\mathrm{d}\zeta}\hat{f}(\zeta) = -\frac{2\pi}{a}\zeta\hat{f}(\zeta).$$

Thus,

$$\frac{\mathrm{d}}{\mathrm{d}\zeta}\left(e^{\frac{\pi\zeta^2}{a}}\hat{f}(\zeta)\right) = e^{(\pi\zeta^2)/a}\left(-\frac{2\pi}{a}\zeta\hat{f}(\zeta)\right) + \frac{2\pi\zeta}{a}e^{(\pi\zeta^2)/a}\hat{f}(\zeta) = 0,$$

and thus $e^{\frac{\pi\zeta^2}{a}}\hat{f}(\zeta)$ is constant in ζ so $e^{\frac{\pi\zeta^2}{a}}\hat{f}(\zeta) = \hat{f}(0) = \int_{-\infty}^{\infty} e^{-\pi a x^2}\,\mathrm{d}x = \frac{1}{\sqrt{a}}$. Thus, $\hat{f}(\zeta) = \frac{1}{\sqrt{a}}e^{-\frac{\pi\zeta^2}{a}}$. ■

With this, we are ready to define the inverse Fourier transform;

↪**Definition 3.3** (Inverse Fourier Transform): If $f \in L^1(\mathbb{R})$, then

$$\check{f}(x) := \int_{\mathbb{R}} f(\zeta) e^{2\pi i \zeta x} \, \mathrm{d}\zeta = \widehat{f(- \cdot)}(x).$$

Remark 3.7: By similar computations to before, $f \in L^1(\mathbb{R})$ implies $\check{f} \in L^\infty(\mathbb{R}) \cap C(\mathbb{R})$.

Remark 3.8: One would hope $\check{\hat{f}} = f$. However, if we check, naively,

$$\check{\hat{f}}(x) = \int \left(\int f(x) e^{-2\pi i \zeta y} \, \mathrm{d}y \right) e^{2\pi i \zeta x} \, \mathrm{d}\zeta;$$

however the integral may not be finite in general, i.e. we cannot switch the integrals for free. We must be more careful, in short.

↪**Theorem 3.4** (Fourier Inversion): If $f \in L^1(\mathbb{R})$ and $\hat{f} \in L^1(\mathbb{R})$, then f agrees almost everywhere with some $f_0 \in C(\mathbb{R})$, and $\check{\hat{f}} = \hat{\check{f}} = f_0$.

Proof. Let $\varepsilon > 0$ and $x \in \mathbb{R}$. Let $\varphi(\zeta) := e^{2\pi i x \zeta} e^{-\pi \varepsilon \zeta^2}$. Then,

$$\begin{aligned}
\hat{\varphi}(y) &= \int \varphi(\zeta) e^{-2\pi i y \zeta} \, \mathrm{d}\zeta \\
&= \int e^{2\pi i x \zeta} e^{-\pi \varepsilon \zeta^2} e^{-2\pi i y \zeta} \, \mathrm{d}\zeta \\
&= \widehat{e^{2\pi i x (\cdot)} e^{-\pi \varepsilon (\cdot)^2}}(y) \\
&= \tau_x \widehat{e^{-\pi \varepsilon (\cdot)^2}}(y) \\
&= \tau_x \left(\frac{1}{\sqrt{\varepsilon}} e^{-\frac{\pi y^2}{\varepsilon}} \right) \\
&= \frac{1}{\sqrt{\varepsilon}} e^{-\frac{\pi}{\varepsilon}(y-x)^2}.
\end{aligned}$$

Since $\int f \hat{\varphi} \, \mathrm{d}y = \int \hat{f} \varphi \, \mathrm{d}y$, we find

$$\int f(y) \frac{1}{\sqrt{\varepsilon}} e^{-\frac{\pi}{\varepsilon}(x-y)^2} = \int \hat{f}(y) \varphi(y) \, \mathrm{d}y.$$

Let $K_\varepsilon(y) := \frac{1}{\sqrt{\varepsilon}} e^{-\frac{\pi}{\varepsilon} y^2}$. Recall that this is the good kernel that we used in the proof of the Weierstrass Approximation Theorem. In particular, the formula above can be written

$$(f * K_\varepsilon)(x) = \int \hat{f}(y) e^{2\pi i x y} e^{-\pi \varepsilon y^2} \, \mathrm{d}y. \qquad \circledast$$

Recall that if f is continuous at x and compactly-supported, then $\lim_{\varepsilon\to 0}|(f * K_\varepsilon)(x) - f(x)| = 0$. This implies that for every $f \in L^1(\mathbb{R})$ $\lim_{\varepsilon \to 0} \|(f * K_\varepsilon) - f\|_1 = 0$, by an approximation argument by $C_c(\mathbb{R})$. So, taking $\varepsilon \to 0$ in ⊛, $f(x) \underset{L^1(\mathbb{R})}{=} \lim_{\varepsilon\to 0} \int \hat{f}(y) e^{2\pi i x y} e^{-\pi \varepsilon y^2} \,\mathrm{d}y$. $\hat{f} \in L^1(\mathbb{R})$, so by DCT we can pass the limit inside, so

$$f(x) \underset{L^1(\mathbb{R})}{=} \int \hat{f}(y) e^{2\pi i x y} \,\mathrm{d}y = \check{\hat{f}}(x).$$

This equality in $L^1(\mathbb{R})$ thus gives $\check{\hat{f}} = f$ a.e.. A similar proof follows for showing $\hat{\check{f}} = f$ a.e. by replacing $e^{2\pi i x}$ with $e^{-2\pi i x}$ everywhere it appears. Since $\hat{\check{f}}, \check{\hat{f}}$ are continuous by our remarks earlier, it follows that f is equal to a continuous function almost everywhere. ■

So far, all we've worked with is $f \in L^1(\mathbb{R})$, which results in $\hat{f} \in L^\infty(\mathbb{R})$. Really, we'd like to extend the Fourier transform to act on $L^2(\mathbb{R})$, since this is a nice Hilbert space. To do so, we need the following:

↪**Theorem 3.5** (Plancherel's Theorem): Let $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Then $\hat{f} \in L^2(\mathbb{R})$ and $\|f\|_{L^2(\mathbb{R})} = \left\|\hat{f}\right\|_{L^2(\mathbb{R})}$.

Remark 3.9: One can view Plancherel's Theorem as a type of continuous analog of Parseval's identity for Fourier Series.

Proof. Let $f(x) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, and put $g(x) := \overline{f(-x)}$, noting that then $g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ as well. Put

$$w(x) := (f * g)(x).$$

By Young's,

$$\|w\|_{L^1(\mathbb{R})} \leq \|f\|_{L^1(\mathbb{R})} \|g\|_{L^1(\mathbb{R})} < \infty$$

so $w \in L^1(\mathbb{R})$.

We claim w continuous at 0. For h sufficiently small, we find

$$\begin{aligned}
|w(h) - w(0)| &= \left| \int_{\mathbb{R}} f(h-y) g(y) \,\mathrm{d}y - \int_{\mathbb{R}} f(-y) g(y) \,\mathrm{d}y \right| \\
&= \left| \int_{\mathbb{R}} (f(h-y) - f(-y)) g(y) \,\mathrm{d}y \right| \\
&\leq \|f(h - \cdot) - f(-\,\cdot)\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})} \\
&\leq \|\tau_h f - f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})}.
\end{aligned}$$

Let $\tilde{f} \in C_c(\mathbb{R})$ such that $\left\|f - \tilde{f}\right\|_{L^2(\mathbb{R})} < \eta$ for some small $\eta > 0$. Then we further bound

$$\begin{aligned}
\|\tau_h f - f\|_{L^2(\mathbb{R})} &\le \left\|\tau_h f - \tau_h \tilde{f}\right\|_{L^2(\mathbb{R})} + \left\|\tau_h \tilde{f} - \tilde{f}\right\|_{L^2(\mathbb{R})} + \left\|\tilde{f} - f\right\|_{L^2(\mathbb{R})} \\
\text{(since norm translation invariant)} \qquad &= 2\left\|\tilde{f} - f\right\|_{L^2(\mathbb{R})} + \left\|\tau_h \tilde{f} - \tilde{f}\right\|_{L^2(\mathbb{R})} \\
&\le 2\eta + \left\|\tau_h \tilde{f} - \tilde{f}\right\|_{L^2(\mathbb{R})}.
\end{aligned}$$

Now, $\tilde{f} \in C_c(\mathbb{R})$ and thus is uniformly continuous hence $\left|\tau_h \tilde{f}(x) - \tilde{f}(x)\right| \to 0$ uniformly on $\mathbb{R}$ hence $\left\|\tau_h \tilde{f} - \tilde{f}\right\|_{L^2(\mathbb{R})} \to 0$ as well, as $h \to 0$. Finally, since $\|g\|_{L^2(\mathbb{R})}$ finite, we conclude indeed w continuous at 0.

Next, notice that $\hat{w} = \hat{f} \cdot \hat{g}$, and

$$\begin{aligned}
\hat{g}(\zeta) = \widehat{\overline{f(-\,\cdot)}}(\zeta) &= \int_{\mathbb{R}} \overline{f(-x)} e^{-2\pi i x \zeta} \, \mathrm{d}x \\
&= \int_{\mathbb{R}} \overline{f(-x) e^{2\pi i x \zeta}} \, \mathrm{d}x \\
&= \overline{\int_{\mathbb{R}} f(-x) e^{2\pi i x \zeta} \, \mathrm{d}x} \\
&= \overline{\int_{\mathbb{R}} f(x) e^{-2\pi i x \zeta} \, \mathrm{d}x} \\
&= \overline{\hat{f}(\zeta)},
\end{aligned}$$

so

$$\hat{w} = \hat{f} \cdot \overline{\hat{f}} = \left|\hat{f}\right|^2 \ge 0.$$

Recall our good kernel from the Weierstrass Approximation Theorem, $K_\varepsilon(y) = \frac{1}{\sqrt{\varepsilon}} e^{-\frac{\pi}{\varepsilon} y^2} = \widehat{e^{-\pi\varepsilon(\cdot)^2}}$. So,

$$\begin{aligned}
\int \hat{w}(y) e^{-\pi\varepsilon y^2} \, \mathrm{d}y &= \int w(y) \frac{1}{\sqrt{\varepsilon}} e^{-(\pi y^2)/\varepsilon} \, \mathrm{d}y \\
&= \int w(y) K_\varepsilon(y) \, \mathrm{d}y \\
\text{(by symmetry)} \qquad &= \int w(y) K_\varepsilon(-y) \, \mathrm{d}y \\
&= (w * K_\varepsilon)(0).
\end{aligned}$$

On the LHS, $\hat{w} \ge 0$ so $\hat{w}(y) e^{-\pi\varepsilon y^2} \underset{\varepsilon\to 0^+}{\nearrow} \hat{w}(y)$ so by monotone convergence, $\int \hat{w} e^{-\pi\varepsilon y^2} \underset{\varepsilon\to 0}{\to} \int \hat{w}(y) \, \mathrm{d}y$. On the other hand, we claim $(w * K_\varepsilon)(0) \underset{\varepsilon\to 0}{\to} w(0)$ (this isn't immediate from the fact that K_ε is a good kernel because we don't know a priori that w (essentially) bounded). Supposing this claim holds, this implies $\int \hat{w}(y) \, \mathrm{d}y = w(0)$, hence

$$\int \hat{w}(y)\,\mathrm{d}y = \int |\hat{f}(y)|^2\,\mathrm{d}y = w(0)$$
$$= (f * g)(0)$$
$$= \int f(y)\overline{f(0-(-y))}\,\mathrm{d}y$$
$$= \int |f(y)|^2\,\mathrm{d}y,$$

which precisely means $\|\hat{f}\|_{L^2(\mathbb{R})} = \|f\|_{L^2(\mathbb{R})}$.

To prove the claim of $(K_\varepsilon * w)(0) \to w(0)$, let $\eta > 0$. Since w continuous at 0, there is a $\delta > 0$ such that $|y| < \delta \Rightarrow |w(y) - w(0)| < \eta$. Then,

$$\left|\int w(0-y)K_\varepsilon(y)\,\mathrm{d}y - w(0)\right| = \left|\int (w(-y) - w(0))K_\varepsilon(y)\,\mathrm{d}y\right|$$
$$\leq \eta \int_{|y|<\delta} K_\varepsilon(y)\,\mathrm{d}y + \int_{|y|>\delta} |w(0)|K_\varepsilon(y)\,\mathrm{d}y + \int_{|y|>\delta} |w(-y)|K_\varepsilon(y)\,\mathrm{d}y$$
$$\leq \eta \cdot 1 + \underbrace{|w(0)|}_{\substack{w \text{ cnts at } 0 \\ \text{so this finite}}} \cdot \underbrace{\int_{|y|>\delta} K_\varepsilon(y)\,\mathrm{d}y}_{\xrightarrow[\varepsilon\to 0]{} 0 \text{ since good kernel}} + \int_{|y|>\delta} |w(-y)|K_\varepsilon(y)\,\mathrm{d}y.$$

It remains to show the last term $\to 0$. We have

$$\int_{|y|>\delta} |w(-y)|\; K_\varepsilon(y)\,\mathrm{d}y \leq \int_{|y|\;>\delta} |w(-y)|\; \frac{1}{\sqrt{\varepsilon}} e^{-\pi\frac{\delta^2}{\varepsilon}}\,\mathrm{d}y$$
$$\leq \underbrace{\frac{1}{\sqrt{\varepsilon}} e^{-\pi\frac{\delta^2}{\varepsilon}}}_{\to 0 \text{ as } \varepsilon\to 0} \cdot \|w\|_{L^1(\mathbb{R})} \to 0.$$

This completes the proof. ■

With these, we can extend the definition of $\hat{f}$ to $f \in L^2(\mathbb{R})$.

Let $f \in L^2(\mathbb{R})$. Then, there are $\{f_k\} \subseteq C_c^\infty(\mathbb{R})$ such that $f_k \to f$ in $L^2(\mathbb{R})$. Since $\{f_k\} \subseteq C_c^\infty(\mathbb{R})$, $f_k \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. So, by Plancherel's,

$$\|f_j - f_k\|_{L^2(\mathbb{R})} = \left\|\widehat{f_j - f_k}\right\|_{L^2(\mathbb{R})} = \left\|\hat{f}_j - \hat{f}_k\right\|_{L^2(\mathbb{R})}.$$

So in particular, $\left\{\hat{f}_k\right\}$ also Cauchy in $L^2(\mathbb{R})$ so by completeness converges. Thus, we simply define the Fourier transform of f as the limit of these, namely,

↪**Definition 3.4** (Fourier Transform on $L^2(\mathbb{R})$): Let $f \in L^2(\mathbb{R})$ and $\{f_k\} \subseteq C_c^\infty(\mathbb{R})$ such that $f_k \to f$ in $L^2(\mathbb{R})$. Then, we define the Fourier transform of f to be

$$\hat{f}(\zeta) := \lim_{j\to\infty} \hat{f}_j(\zeta),$$

with the limit taken in $L^2(\mathbb{R})$.

It's not obvious that this is well-defined a priori. Let f_k, f_k' be two sequences in $C_c^\infty(\mathbb{R})$ converging to f in $L^2(\mathbb{R})$, and suppose $\hat{f}_k \to \hat{f}, \hat{f}_k' \to \hat{f}'$ in $L^2(\mathbb{R})$. We need to show that $\hat{f} = \hat{f}'$ in $L^2(\mathbb{R})$. Since $f_k, f_k' \to f$ in $L^2(\mathbb{R})$,

$$\|f_k - f_k'\|_2 \to 0,$$

so also by Plancherel's,

$$\left\|\hat{f}_k - \hat{f'}_k\right\|_2 \to \left\|\hat{f} - \hat{f}'\right\|_2 = 0.$$

Denote by $C_0(\mathbb{R}) := \left\{f \in C(\mathbb{R}) \mid \lim_{|x| \to \infty} |f(x)| = 0\right\}$.

↪**Proposition 3.4**:
1. If $xf(x) \in L^1(\mathbb{R})$, $\hat{f} \in C^1(\mathbb{R})$ and $\partial_\zeta \hat{f}(\zeta) = (-2\widehat{\pi i(\cdot))f}(\cdot)(\zeta)$
2. If $f \in C^1(\mathbb{R}) \cap C_0(\mathbb{R})$ and $\partial_x f \in L^1(\mathbb{R})$, then $\widehat{\partial_x f}(\zeta) = (2\pi i\zeta)\hat{f}(\zeta)$
3. If $f \in L^1(\mathbb{R})$, then $\hat{f} \in C_0(\mathbb{R})$ ("Riemann-Lebesgue" type result)

Proof. We prove only 3. If $f \in L^1(\mathbb{R})$, let $\{g_n\} \subseteq C^1(\mathbb{R}) \cap C_c(\mathbb{R})$ such that $g_n \to f$ in $L^1(\mathbb{R})$. Then, g_n' are compactly supported and continuous so $g_n' \in L^1(\mathbb{R})$. Thus, $\widehat{g_n'} \in L^\infty(\mathbb{R})$. By 2., $\widehat{g_n'}(\zeta) = (2\pi i\zeta)\widehat{g_n}(\zeta) \in L^\infty(\mathbb{R})$. Thus is only possible if $\hat{g}_n \in C_0(\mathbb{R})$.

Since $\|g_n - f\|_1 \to 0$,

$$\left\|\widehat{g_n} - \hat{f}\right\|_\infty = \sup_\zeta \left| \int (g_n(x) - f(x)) e^{-2\pi i\zeta x} \, \mathrm{d}x \right| \leq \|g_n - f\|_1 \to 0,$$

so $\widehat{g_n} \to \hat{f}$ in L^∞. Finally, for any n,

$$\lim_{|\zeta| \to \infty} \left|\hat{f}(\zeta)\right| \leq \underbrace{\lim_{|\zeta| \to \infty} |\hat{g}_n(\zeta)|}_{=0} + \left\|\hat{f} - \hat{g}_n\right\|_\infty.$$

Sending then $n \to \infty$, we know that $\left\|\hat{f} - \hat{g}_n\right\|_\infty \to 0$, completing the proof. ■

Remark 3.10: Properties 1., 2. here can be extended to $f \in L^2(\mathbb{R})$ and $\partial_x f \in L^2(\mathbb{R})$, but require more delicate mollifying arguments. 3., however, does not extend.

Remark 3.11: Why is it important to extend $\hat{f}(\zeta)$ to $f \in L^2(\zeta)$? One reason is the analysis of Sobolev Spaces.

The final topic we'll cover is how we can relate Fourier Series to the Fourier Transform.

↪**Theorem 3.6**: If $f \in L^1(\mathbb{R})$, then there exists $Pf : \mathbb{T} \to \mathbb{C}$ defined by

$$Pf(x) = \sum_{k\in\mathbb{Z}} \tau_k f(x) = \sum_{k\in\mathbb{Z}} f(x-k)$$

(that is, we tacitly claim this summation converges pointwise a.e. and in $L^1(\mathbb{T})$), and

$$\|Pf\|_{L^1(\mathbb{T})} \le \|f\|_{L^1(\mathbb{R})}.$$

Also, for every $k \in \mathbb{Z}$,

$$\hat{f}(k) = \widehat{Pf}(k),$$

where the first $\hat{\cdot}$ is the Fourier transform on $\mathbb{R}$ and the second on $\mathbb{T}$.

PROOF. Let $Q = \left[-\frac{1}{2}, \frac{1}{2}\right)$ so $\mathbb{R} = \bigsqcup_{j\in\mathbb{Z}} Q + j$. Then,

$$\begin{aligned}
\int_{\mathbb{R}} |f(x)| \,\mathrm{d}x &= \sum_{j\in\mathbb{Z}} \int_{Q+j} |f(x)| \,\mathrm{d}x \\
&= \sum_{j\in\mathbb{Z}} \int_{Q} \underbrace{|f(x-j)|}_{\ge 0} \,\mathrm{d}x \\
\text{(Tonelli's)} \qquad &= \int_Q \sum_{j\in\mathbb{Z}} |f(x-j)| \,\mathrm{d}x.
\end{aligned}$$

Thus,

$$\int_Q \sum_{j\in\mathbb{Z}} \tau_j f(x) \,\mathrm{d}x \le \int_Q \sum_{j\in\mathbb{Z}} |f(x-j)| \,\mathrm{d}x = \|f\|_{L^1(\mathbb{R})}.$$

So, Pf as defined above has

$$\|Pf\|_{L^1(\mathbb{T})} \le \|f\|_{L^1(\mathbb{R})},$$

and also Pf is finite a.e.. Hence, the sum in question defining $Pf(x)$ converges a.e.. Moreover,

$$\begin{aligned}
\widehat{Pf}(k) &= \underbrace{\int_Q \sum_{j\in\mathbb{Z}} f(x-j) e^{-2\pi i k x}}_{\in L^1(Q)} \,\mathrm{d}x \\
\text{(By Fubini)} \qquad &= \sum_{j\in\mathbb{Z}} \int_Q f(x-j) e^{-2\pi i k x} \,\mathrm{d}x \\
&= \sum_{j\in\mathbb{Z}} \int_{Q-j} f(x) \underbrace{e^{-2\pi i k(x+j)}}_{\substack{= e^{-2\pi i k x} \\ \text{since } e^{-2\pi i k j} = 1}} \,\mathrm{d}x \\
&= \int_{\mathbb{R}} f(x) e^{-2\pi i k x} \,\mathrm{d}x = \hat{f}(k).
\end{aligned}$$

■

This series Pf is called the *periodization* of f.

↪**Theorem 3.7** (Poisson Summation Formula): Let $f \in C(\mathbb{R})$ such that there are $C, \varepsilon > 0$ such that $|f(x)| \leq C(1+|x|)^{-(1+\varepsilon)}$ (so namely $f \in L^1(\mathbb{R})$) and similarly $\left|\hat{f}(\zeta)\right| \leq C(1+|\zeta|)^{-(1+\varepsilon)}$. Then,

$$\sum_{k\in\mathbb{Z}} f(x+k) = \sum_{k\in\mathbb{Z}} \hat{f}(k)e^{2\pi ikx},$$

where both series converge absolutely and uniformly on $\mathbb{T}$. In particular, if $x = 0$,

$$\sum_{k\in\mathbb{Z}} f(k) = \sum_{k\in\mathbb{Z}} \hat{f}(k).$$

Remark 3.12: By the last remark, $\hat{f}(k) = \widehat{Pf}(k)$. So, this theorem says "periodized f" = Pf = "Fourier series of Pf".

Proof. Fix $x \in \mathbb{R}$ then

$$\begin{aligned}
\left|\sum_{k\in\mathbb{Z}} f(x+k)\right| &\leq \sum_{k\in\mathbb{Z}} |f(x+k)| \\
&\leq \int_{\mathbb{R}} |f(x+y)| \,\mathrm{d}y \\
&\leq \int_{\mathbb{R}} \frac{C}{(1+|x+y|)^{1+\varepsilon}} \,\mathrm{d}y \\
&= \int_{\mathbb{R}} \frac{C}{(1+|y|)^{1+\varepsilon}} \,\mathrm{d}y \\
&= -\frac{C}{(1+|y|)^{\varepsilon}}\Bigg|_{y=-\infty}^{\infty} \leq C,
\end{aligned}$$

hence the series absolutely converges, and since our bound is independent of x, it also converges uniformly. Since $S_N(x) := \sum_{k=-N}^{N} f(x+k)$ is continuous for each N and $S_N \to Pf$ uniformly, Pf itself is continuous, in $C(\mathbb{T})$ so thus also in $L^2(\mathbb{T})$. Thus, by Hilbert space theory,

$$Pf(x) = \sum_{k\in\mathbb{Z}} \widehat{Pf}(k)e^{2\pi ikx},$$

in $L^2(\mathbb{T})$. By the last result, $\widehat{Pf}(k) = \hat{f}(k)$, thus

$$Pf(x) = \sum_{k\in\mathbb{Z}} \hat{f}(k)e^{2\pi ikx}.$$

Finally, by the same computation as before, $\sum_{k\in\mathbb{Z}} \hat{f}(k)e^{2\pi ikx}$ will also converge absolutely and uniformly as well, call it $\tilde{P}f(x)$. Thus, we claim $\tilde{P}f = Pf$. Indeed, Pf is continuous, and $\tilde{P}f = \lim_{N\to\infty} \sum_{k=-N}^{N} \hat{f}(k)e^{2\pi ikx}$ so $\tilde{P}f$ also continuous. So,

$$\hat{S}_N(x) \underset{L^2(\mathbb{T})}{\rightarrow} Pf(x)$$

$$\hat{S}_N(x) \underset{\text{uniform}}{\rightarrow} \tilde{P}f(x),$$

and $Pf, \tilde{P}f$ are both continuous hence $Pf \equiv \tilde{P}f$. Thus, indeed $Pf = \sum \hat{f}(k)e^{2\pi ikx}$ as we aimed to show. ■