# 1 Lebesgue Integration Theory

## 1.1 Review of measure theory

**Definition.** Given a set $E$, a *$\sigma$-algebra* on $E$ is a collection $\mathcal{E}$ of subsets of $E$ such that:

(i) $E \in \mathcal{E}$;

(ii) $A \in \mathcal{E} \Rightarrow A^c = E \setminus A \in \mathcal{A}$;

(iii) $A_n \in \mathcal{E}$, $n \in \mathbb{N} \Rightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{E}$.

$(E, \mathcal{E})$ is called a *measurable space*, and any $A \in \mathcal{E}$ is called a *measurable set.*

Given a collection $\mathcal{A}$ of subsets of $E$, $\sigma(\mathcal{A})$ is the smallest $\sigma$-algebra containing $\mathcal{A}$.

**Definition.** A *measure* on $(E, \mathcal{E})$ is a function $\mu : \mathcal{E} \to [0, \infty]$ such that

(i) $\mu(\emptyset) = 0$;

(ii) $A_n \in \mathcal{E}$, $n \in \mathbb{N}$ disjoint $\Rightarrow \mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n)$.

$(E, \mathcal{E}, \mu)$ is called a *measure space.*

**Definition** (Borel measure)**.** If $(E, \tau)$ is a topological space, then $\sigma(\tau)$ is called a *Borel algebra*, denoted $\mathcal{B}(E)$, and a measure on $(E, \mathcal{B}(E))$ is called a *Borel measure.*

**Example.** $E = \mathbb{R}^n$, $\mu$ the Lebesgue measure satisfying $\mu((a_1, b_1) \times \ldots \times (a_n, b_n)) = (b_1 - a_1) \ldots (b_n - a_n)$.

**Notation**: we write $\mu(\mathrm{d}x) = \mathrm{d}x$ and $\mu(A) = |A|$ when $\mu$ is the Lebesgue measure.

**Definition** (Measurable function)**.** Let $(E, \mathcal{E})$ and $(F, \mathcal{F})$ be measurable spaces. Then $f : E \to F$ is *measurable* if $f^{-1}(A) \in \mathcal{E}$ whenever $A \in \mathcal{F}$. If $(E, \mathcal{E})$ and $(F, \mathcal{F})$ are Borel algebras, a measurable function is called a *Borel function.* Special case: $(F, \mathcal{F}) = ([0, \infty], \mathcal{B}([0, \infty]))$, then $f : E \to F$ is called a *non-negative measurable function.*

**Fact.** The class of measurable functions is closed under addition, multiplication and taking (pointwise) limits.

**Definition.** $f : E \to F$ ($F = [0, \infty]$ or $\mathbb{R}^n$ or $\mathbb{C}^n$) is a *simple function* if $f = \sum_{k=1}^{K} a_k \mathbb{1}_{A_k}$ for some $K \in \mathbb{N}$, $a_k \in F$, $A_k \in \mathcal{E}$. For a simple function, the integral is

$$\int f \mathrm{d}\mu = \int f(x) \mathrm{d}\mu(x) = \sum_{k=1}^{K} a_k \mu(A_k) \ (0 \cdot \infty := 0).$$

For a non-negative measurable $f$, we define

$$\int f \mathrm{d}\mu = \sup \left\{ \int g \mathrm{d}\mu : g \text{ simple }, 0 \le g \le f \right\}.$$

**Definition.** A measurable function $f : E \to \mathbb{R}$ is said to be *integrable* if $\int |f| \mathrm{d}\mu < \infty$. Write $f = f_+ - f_-$ with $f_\pm$ non-negative, measurable, $\int f_\pm \mathrm{d}\mu < \infty$, and then $\int f \mathrm{d}\mu = \int f_+ \mathrm{d}\mu - \int f_- \mathrm{d}\mu$. For $f : E \to \mathbb{R}^n$, this is applied in each component.

**Theorem** (Monotone convergence theorem)**.** *Let $(E, \mathcal{E}, \mu)$ be a measure space, and let $(f_n)$ be a (pointwise) increasing sequence of non-negative functions on $E$ converging to $f$. Then*

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

**Theorem** (Dominated convergence theorem)**.** *Let $(f_n)$ be a sequence of measurable functions on a measure space $(E, \mathcal{E}, \mu)$ such that:*

*(i) $f_n \to f$ pointwise almost everywhere;*

*(ii) $|f_n| \leq g$ almost everywhere for some integrable $g$.*

*Then*

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

## 1.2 $L^p$ spaces

**Definition.** Let $(E, \mathcal{E}, \mu)$ be a measure space. For $p \in [1, \infty)$ and $f : E \to \mathbb{R}$ define

$$||f||_{L^p} = \left( \int_E |f|^p \mathrm{d}\mu \right)^{1/p}$$

and

$$||f||_{L^\infty} = \mathrm{esssup}|f| = \inf\{K : |f| \leq K \text{ a.e}\}.$$

The space $L^p$, $p \in [1, \infty]$ is defined by

$$L^p = L^p(E, \mathcal{E}, \mu) = \{f : E \to \mathbb{R} \text{ measurable} : ||f||_{L^p} < \infty\}/\sim.$$

Where $f \sim g$ if $f = g$ a.e.

**Theorem** (Riesz-Fisher theorem)**.** *$L^p$ is a Banach space for all $p \in [1, \infty]$.*

**Notation**: when $E = \mathbb{R}^n$, $\mu$ the Lebesgue measure, write $L^p(E, \mu) = L^p(\mathbb{R}^n)$.

**Fact.** For $p \in [1, \infty)$, the simple functions $f$ with $\mu(\{x : f(x) \neq 0\}) < \infty$ are dense in $L^p$. For $p = \infty$ we can drop the condition on the measure of the support.

**Definition.** For $f, g : \mathbb{R}^n \to \mathbb{R}$, the *convolution* $f * g$ is defined by

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

provided the integral exists. Note that $f * g = g * f$, convolution is associative, and $\mu(f * g) = \mu(f)\mu(g)$.

**Theorem.** *$C_c^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for $p \in [1, \infty)$.*

Before we prove the theorem, we will need some preliminary results.

**Remark.** This theorem is false for $p = \infty$.

**Notation**: a multiindex is $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}^n_{\geq 0}$. Set $|\alpha| = \alpha_1 + \ldots + \alpha_n$; $\alpha! = \alpha_1! \ldots \alpha_n!$; $X^\alpha = X_1^{\alpha_1} \ldots X_n^{\alpha_n}$ for $X \in \mathbb{R}^n$; $\nabla^\alpha f = D^\alpha f = \frac{\partial^\alpha f}{\partial x^\alpha} = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}}$.

**Definition.** We say $f \in L^p_{\text{loc}}(\mathbb{R}^n)$ if $f \mathbb{1}_K \in L^p(\mathbb{R}^n)$ for any $K \subseteq \mathbb{R}^n$ compact.

**Proposition.** Let $f \in L^1_{\text{loc}}(\mathbb{R}^n)$, $g \in C^k_c(\mathbb{R}^n)$, some $k \geq 0$. Then $f * g \in C^k(\mathbb{R}^n)$ and $\nabla^\alpha (f * g) = f * (\nabla^\alpha g)$ for all $|\alpha| \leq k$.

*Proof.* First we check for $k = 0$. Set $T_z f(x) = f(x - z)$, $z \in \mathbb{R}^n$. Then $T_z(f * g) = f * (T_z g)$. Also $T_z g(x) \to g(x)$ for all $x$ as $z \to 0$ (continuity of $g$). Furthermore $|T_z g(x)| \leq ||g||_{L^\infty} \mathbb{1}_{B_R(0)}(x)$ if $|x| + 1 \leq R$, $|z| < 1$ (we can just take $R$ large enough so it holds everywhere since $g$ has compact support). Then $|f(y) T_z g(x - y)| \leq C |f(y)| \mathbb{1}_{B_R(0)}(x - y)$, for $C := ||g||_{L^\infty}$.

Since $f \in L^1_{\text{loc}}(\mathbb{R}^n)$, $|f(y)| \mathbb{1}_{B_R(0)}(x - y)$ is integrable in $y$, so by the dominated convergence theorem,

$$T_z(f * g) = (f * T_z g)(x) = \int_{\mathbb{R}^n} f(y) T_z g(x - y) \mathrm{d}y \xrightarrow{z \to 0} \int_{\mathbb{R}^n} f(y) g(x - y) \mathrm{d}y = (f * g)(x).$$

And so $f * g \in C^0$. Now let $k = 1$. Let $\nabla_i^h g(x) = \frac{g(x + h e_i) - g(x)}{h}$, where $e_i$ is the $i$th unit vector. Then $\nabla_i^h g(x) \to \nabla_i g(x)$ as $h \to 0$.

By the mean value theorem, there exists $t \in [-h, h]$ such that

$$\nabla_i^h g(x) = \nabla_i g(x + t e_i) \Rightarrow |\nabla_i^h g(x)| \leq ||\nabla_i g||_{L^\infty} \mathbb{1}_{B_R(0)}(x).$$

Again by the dominated convergence theorem, $\nabla_i^h (f * g) = f * (\nabla_i^h g) \to f * \nabla_i g$. Thus $f * g \in C^1$. The case $k > 1$ is similar, with induction. □

**Proposition** (Minkowski's integral inequality)**.** Let $p \in [1, \infty)$ and $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ Borel. Then

$$\left[ \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} f(x, y) \mathrm{d}x \right|^p \mathrm{d}y \right]^{1/p} \leq \int_{\mathbb{R}^n} \left[ \int_{\mathbb{R}^n} |f(x, y)|^p \mathrm{d}y \right]^{1/p} \mathrm{d}x.$$

*Proof.* Example sheet 1. □

**Proposition.** Let $p \in [1, \infty)$, $g \in L^p(\mathbb{R}^n)$. Then

$$||T_z g - g||_{L^p} \to 0 \text{ as } |z| \to 0.$$

**Remark.** This is not true for $p = \infty$. Let $\theta(x) = \mathbb{1}_{x \geq 0}$. Then $||T_z \theta - \theta||_{L^\infty} = 1$ if $z \neq 0$.

*Proof.* Consider first $g = \mathbb{1}_R$, $R$ a rectangle. Then the result is clear. Hence it also follows for any finite union of rectangles. If $B$ is a Borel set, $|B| < \infty$, then for every $\varepsilon > 0$, there exists a finite union of rectangles $R$ such that

$$||\mathbb{1}_B - \mathbb{1}_R||_{L^p} = |B \triangle R|^{1/p} < \varepsilon.$$

Hence

$$||T_z\mathbb{1}_B - \mathbb{1}_B||_{L^p} \leq \underbrace{||T_z\mathbb{1}_B - T_z\mathbb{1}_R||_{L^p}}_{=||\mathbb{1}_B - \mathbb{1}_R||_{L^p} < \varepsilon} + \underbrace{||T_z\mathbb{1}_R - \mathbb{1}_R||_{L^p}}_{<\varepsilon \text{ for } |z| \text{ small}} + \underbrace{||\mathbb{1}_R - \mathbb{1}_B||_{L^p}}_{<\varepsilon}.$$

Thus the result holds for $g = \mathbb{1}_B$, $B \in \mathcal{B}(\mathbb{R}^n)$. Thus the result holds for simple functions $g$. Finally, for any $g \in L^p$, there is a $\tilde{g}$ simple such that $||g - \tilde{g}||_{L^p} < \varepsilon$. Then

$$||T_z g - g||_{L^p} \leq \underbrace{||T_z g - T_z\tilde{g}||_{L^p}}_{=||g-\tilde{g}||_{L^p} < \varepsilon} + \underbrace{||T_z\tilde{g} - \tilde{g}||_{L^p}}_{<\varepsilon \text{ for } |z| \text{ small}} + \underbrace{||g - \tilde{g}||_{L^p}}_{<\varepsilon}.$$

□

**Theorem.** *Let $\varphi \in C_c^\infty(\mathbb{R}^n)$ be such that $\varphi \geq 0$ and $\int_{\mathbb{R}^n} \varphi \mathrm{d}x = 1$ and set $\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi(x/\varepsilon)$. Then for any $g \in L^p$, $p \in [1, \infty)$, it follows that $\varphi_\varepsilon * g \in C^\infty(\mathbb{R}^n)$ and $\varphi_\varepsilon * g \to g$ in $L^p$.*

*Proof.* We have

$$\begin{aligned} \varphi_\varepsilon * g(x) - g(x)| &= \left| \int_{\mathbb{R}^n} [\varphi_\varepsilon(y) g(x-y) - g(x)] \, \mathrm{d}y \right| \\ &\underbrace{=}_{z:=y/\varepsilon} \left| \int_{\mathbb{R}^n} \varphi(z) \, [g(x - \varepsilon z) - g(x)] \, \mathrm{d}z \right| \\ &\leq \int_{\mathbb{R}^n} \varphi(z) \, |T_{\varepsilon z} g(x) - g(x)| \, \mathrm{d}z. \end{aligned}$$

Hence

$$\begin{aligned} \|\varphi_\varepsilon * g - g\|_{L^p} &= \left( \int_{\mathbb{R}^n} \underbrace{|\varphi_\varepsilon * g - g|^p}_{\int_{\mathbb{R}^n} \varphi(z) |T_{\varepsilon z} g - g| \mathrm{d}z} \, \mathrm{d}x \right)^{1/p} \\ &\leq \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} \varphi(z)^p |T_{\varepsilon z} g(x) - g(x)|^p \mathrm{d}x \right)^{1/p} \mathrm{d}z \\ &= \int_{\mathbb{R}^n} \varphi(z) \underbrace{||T_{\varepsilon z} g - g||_{L^p}}_{\to 0 \text{ as } \varepsilon \to 0} \mathrm{d}z \end{aligned}$$

where the inequality comes from Minkowski's inequality. Also the final term tends to 0 as $\varepsilon \to 0$ by the DCT since $\varphi(z)||T_{\varepsilon z} g - g||_{L^p}|| \leq 2\varphi(z)||g||_{L^p}$ and $\varphi$ is integrable. □

**Definition.** $\varphi$ as above is called a (smooth) mollifier.

**Corollary.** *$C_c^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$, $p \in [1, \infty)$.*

*Proof.* The previous theorem implies $C^\infty(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$ is dense in $L^p$. Since $||f - f\mathbb{1}_{B_R(0)}||_{L^p} \to 0$ as $R \to \infty$ by the DCT, for $f \in L^p$, applying the theorem with $g = f\mathbb{1}_{B_R(0)}$ it follows that $C_c^\infty(\mathbb{R}^n)$ is dense in $L^p$. □

## 1.3 Lebesgue Differentiation Theorem

Recall:

**Theorem** (Fundamental Theorem of Calculus)**.** *For $f : \mathbb{R} \to \mathbb{R}$ continuous, $F(x) := \int_0^x f(t)\mathrm{d}t$ is differentiable with $F'(x) = f(x)$.*

We actually have a stronger result:

**Theorem** (Lebesgue Differentiation Theorem)**.** *For $f : \mathbb{R}^n \to \mathbb{R}$ integrable,*

$$\lim_{r \to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)| \mathrm{d}y = 0 \text{ for almost all } x.$$

*The $x$ for which this holds are called Lebesgue points.*

We will need a few preliminary results and definitions before we can prove this.

**Corollary.** *If $g \in L^1(\mathbb{R})$ and $G(x) = \int_{-\infty}^x g(t)\mathrm{d}t$, then $G$ is differentiable for almost every $x$ with $G'(x) = g(x)$.*

**Corollary.** *If $\varphi$ is a smooth mollifier and $g \in L^p(\mathbb{R}^n)$, then $\varphi_\varepsilon * g \xrightarrow{\varepsilon \to 0} g$ almost everywhere.*

**Definition.** For $f : \mathbb{R}^n \to \mathbb{R}$ integrable, the *Hardy-Littlewood Maximal Function* $\mathrm{Mf} : \mathbb{R}^n \to [0, \infty]$ is defined by

$$\mathrm{Mf}(x) = \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)|\mathrm{d}y.$$

**Remark.** We sometimes write $⨍_{B_r(x)} |f(y)|\mathrm{d}y$ for $\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)|\mathrm{d}y$.

**Lemma** (Wiener's covering lemma)**.** *If $K$ is compact and $K \subseteq \bigcup_{i=1}^N B_i$ for open balls $(B_i)_{i=1}^N$, there exists a subcollection $(B_{i_k})_k$ of disjoint balls such that*

$$\left| \bigcup_{i=1}^N B_i \right| \leq 3^n \sum_k |B_{i_k}|.$$

*Proof.* Example sheet. □

**Proposition.** Take $f \in L^1(\mathbb{R}^n)$. Then Mf is a Borel function, finite almost everywhere, and

$$|\underbrace{\{\mathrm{Mf} > \lambda\}}_{:= A_\lambda}| \leq \frac{3^n}{\lambda} ||f||_{L^1}.$$

*Proof.* For each $x \in A_\lambda$, there exists $r_x > 0$ such that

$$\frac{1}{B_{r_x}(x)} \int_{B_{r_x}(x)} |f(y)|\mathrm{d}y > \lambda.$$

We claim that $A_\lambda$ is open. Then we will have shown Mf is Borel as the $A_\lambda = (\mathrm{Mf})^{-1}((\lambda, \infty])$ are open, and the sets $(\lambda, \infty]$ generate the Borel $\sigma$-algebra.

We'll actually show $A_\lambda^c$ is closed. Suppose $(x_k)_{k \geq 1}$ is a sequence in $A_\lambda^c$ with $x_k \to x$. Suppose $x \in A_\lambda$. By the Dominated Convergence Theorem,

$$\frac{1}{B_{r_x}(x_k)} \int_{B_{r_x}(x_k)} |f(y)|\mathrm{d}y \to \frac{1}{B_{r_x}(x)} \int_{B_{r_x}(x)} |f(y)|\mathrm{d}y.$$

Since $x_k \notin A_\lambda$, the LHS is $\leq \lambda$ for all $k$, but the RHS is $> \lambda$ which is impossible. Hence $x \in A_\lambda^c$ and $A_\lambda^c$ is closed.

To prove the inequality, let $K \subseteq A_\lambda$ be compact. Since $\{B_{r_x}(x)\}_{x \in A_\lambda}$ is an open cover of $K$, there exists a finite subcover $K \subseteq \bigcup_{i=1}^N B_i$, where $B_i = B_{r_x}(x)$ for some $x \in A_\lambda$. Now take a subcollection $(B_{i_k})_k$ of disjoint balls as in Wiener's covering lemma.

Since $\frac{1}{|B_i|}\int_{B_i}|f(y)|\mathrm{d}y > \lambda$, it follows that $|B_i| < \frac{1}{\lambda}\int_{B_i}|f(y)|\mathrm{d}y$. Hence

$$|K| \leq 3^n \sum_k |B_{i_k}| < \frac{3^n}{\lambda}\sum_k \int_{B_{i_k}} |f(y)|\mathrm{d}y \leq \frac{3^n}{\lambda}\int_{\mathbb{R}^n} |f(y)|\mathrm{d}y.$$

Since this holds for any $K \subseteq A_\lambda$ compact, by regularity of the Lebesgue measure, it also holds for $A_\lambda$. In particular, $|\{\mathrm{Mf} = \infty\}| \leq |\{\mathrm{Mf} > \lambda\}| \xrightarrow{\lambda\to\infty} 0$, i,e $\mathrm{Mf} < \infty$ almost everywhere. $\square$

Now we are ready to prove:

**Theorem** (Lebesgue Differentiation Theorem)**.** *For $f : \mathbb{R}^n \to \mathbb{R}$ integrable,*

$$\lim_{r\to 0} \frac{1}{|B_r(x)|}\int_{B_r(x)} |f(y)-f(x)|\mathrm{d}y = 0 \text{ for almost all } x.$$

*The $x$ for which this holds are called Lebesgue points.*

*Proof.* Let

$$A_\lambda = \left\{ x \in \mathbb{R}^n : \limsup_{r\to 0} ⨍_{B_r(x)} |f(y)-f(x)|\mathrm{d}y > 2\lambda \right\}$$

Then it suffices to show $|A_\lambda| = 0$ for any $\lambda > 0$. Indeed, the non-Lebesgue points are then $\bigcup_n A_{1/n}$, a countable union of sets of measure 0.

Given $\varepsilon > 0$, let $g \in C_c^\infty(\mathbb{R}^n)$ be such that $||f-g||_{L^1} < \varepsilon$. Then

$$⨍_{B_r(x)} |f(y)-f(x)|\mathrm{d}y \le \underbrace{⨍_{B_r(x)} |f(y)-g(y)|\mathrm{d}y}_{\le M(f-g)(x)} + |f(x)-g(x)| + \underbrace{⨍_{B_r(x)} |g(x)-g(y)|\mathrm{d}y}_{\to 0 \text{ since } g\in C^\infty}.$$

$$\implies \limsup_{r\to 0} ⨍_{B_r(x)} |f(y)-f(x)|\mathrm{d}y \le M(f-g)(x) + |f(x)-g(x)|.$$

If $x \in A_\lambda$, then either $M(f-g)(x) > \lambda$ or $|f(x)-g(x)| > \lambda$. The Hardy-Littlewood maximal inequality says $|\{M(f-g) > \lambda\}| \le \frac{3^n}{\lambda}||f-g||_{L^1}$. Then by Markov's inequality $|\{|f-g| > \lambda\}| \le \frac{1}{\lambda}||f-g||_{L^1}$. Hence

$$|A_\lambda| \le \frac{3^n+1}{\lambda}||f-g||_{L^1} < \frac{3^{n+1}+1}{\lambda}\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, $|A_\lambda| = 0$. $\square$

## 1.4 Littlewood's Principles

**Theorem** (Egorov)**.** *Let $E \subseteq \mathbb{R}^n$, $|E| < \infty$, and $f_k : E \to \mathbb{R}$, $k \ge 1$ be a sequence of measurable functions such that $f_k \to f$ almost everwhere. Then for every $\varepsilon > 0$, there is a closed subset $A_\varepsilon \subseteq E$ such that $|E \setminus A_\varepsilon| < \varepsilon$ and $f_k \to f$ uniformly on $A_\varepsilon$.*

*Proof.* Without loss of generality, $f_k(x) \to f(x)$ for all $x \in E$ (otherwise restrict to a subset of $E$ of full measure). Let

$$E_k^n = \left\{ x \in E : |f_j(x) - f(x)| < \frac{1}{n} \ \forall j > k \right\}.$$

Then $E^n_{k+1} \supseteq E^n_k$, $\bigcup_k E^n_k = E$, hence $|E^n_k| \uparrow |E|$ as $k \to \infty$. Let $k_n$ be such that $|E \setminus E^n_{k_n}| < 2^{-n}$ and for $N \in \mathbb{N}$ set

$$A_N = \bigcap_{n \geq N} E^n_{k_n} \implies |E \setminus A_N| \leq \sum_{n \geq N} |E \setminus E^n_{k_n}| \leq 2^{-N+1} < \varepsilon \text{ for } N = N_\varepsilon.$$

Now it suffices to show $f_j \to f$ uniformly on $A_N$. Indeed, for $x \in A_N$ and any $n \geq N$, $|f_j(x) - f(x)| < \frac{1}{n}$ for all $j > k_n$. Hence $\limsup_{j\to\infty} \sup_{A_N} |f_j - f| \leq \frac{1}{n}$ for all $n \geq N$, hence $\lim_{j\to\infty} \sup_{A_N} |f_j - f| = 0$. □

**Theorem** (Lusin)**.** *Let $f : E \to \mathbb{R}$ be a Borel function, where $E \subseteq \mathbb{R}^n$ and $|E| < \infty$. Then for every $\varepsilon > 0$, there exists $F_\varepsilon \subseteq E$ closed such that $|E \setminus F_\varepsilon| < \varepsilon$ and $f|_{F_\varepsilon}$ is continuous.*

**Remark.** Careful: this does <u>not</u> mean that $f$ is continuous at $x \in F_\varepsilon$ in the topology of $\mathbb{R}^n$.

*Proof.* First we show that the statement holds for simple functions $f$. Let $f = \sum_{m=1}^M a_m \mathbb{1}_{A_m}$ with the $A_m$ disjoint and $\bigcup_m A_m = E$. Then there are compact sets $K_m \subseteq A_m$ with $|A_m \setminus K_m| < \frac{\varepsilon}{M}$ by regularity of the Lebesgue measure. Then if $F_\varepsilon = \bigcup_m K_m$, $|E \setminus F_\varepsilon| < \varepsilon$. Since $f$ is constant on each $K_m$, and the distance between $K_m$ and $K_{m'}$ is strictly positive for $m \neq m'$ (compactness), this implies $f|_{F_\varepsilon}$ is continuous.

Now we show the statement holds for any measurable $f$. Let $f_n$ be simple functions such that $f_n \to f$ almost everywhere, and $C_n \subseteq E$ be such that $|C_n| < 2^{-n}$ and $f_n|_{E \setminus C_n}$ is continuous for all $n$. By Egorov's Theorem, there exists $A_\varepsilon$ such that $f_n \to f$ uniformly on $A_\varepsilon$ and $|E \setminus A_\varepsilon| < \varepsilon$. Set $F'_\varepsilon = A_\varepsilon \setminus \bigcup_{n \geq N} C_n$ so $|E \setminus F'_\varepsilon| < 2\varepsilon$ for $N = N_\varepsilon$ sufficiently large. Since $f_n|_{F'_\varepsilon}$, $n \geq N$ is continuous $f_n \to f$ uniformly on $F'_\varepsilon$, $f|_{F'_\varepsilon}$ is continuous.

By regularity of the Lebesgue measure, there exists $F_\varepsilon \subseteq F'_\varepsilon$ closed with $|F'_\varepsilon \setminus F_\varepsilon| < \varepsilon$ so $|E \setminus F_\varepsilon| < 3\varepsilon$ and we are done. □

# 2 Banach and Hilbert space analysis

## 2.1 The Hilbert space $L^2$

For any measure space $(E, \mathcal{E}, \mu)$, $L^2(E, \mu)$ is a Hilbert space with inner product

$$\langle f, g \rangle_{L^2} = \int_E \overline{f} g \mathrm{d}\mu.$$

**Definition.** A subset $S = \{u_j\}_{j \in J} \subseteq H$ of a Hilbert space $H$ is

- *Orthogonal* if $\langle u_j, u_k \rangle = 0$ for all $j \neq k$;
- *Orthonormal* if it is orthogonal and $\langle u_j, u_j \rangle = 1$ for all $j$;
- *Complete* if $\overline{\operatorname{span}\{u_j\}} = H$.

A complete orthonormal set is called a *Hilbert basis*.

**Fact.** A Hilbert space is separable (i.e there is a countable dense subset) if and only if there is a countable orthonormal (Hilbert) basis.

**Examples.**

(i) $L^2([-\pi, \pi]), S = \left\{ \frac{1}{\sqrt{2\pi}} e^{-inx} \right\}_{n \in \mathbb{Z}}$. Then $S$ is a Hilbert basis; the Fourier basis (completeness follows from the Stone-Weierstrass theorem & density of $C^\infty$).

(ii) $L^2(\mathbb{R})$, $S = \{\psi_{n,k}\}_{n,k \in \mathbb{Z}}$ where

$$\psi_{n,k}(x) = 2^{n/2} \psi(2^n x - k),$$

$$\psi(x) = \begin{cases} 1 & x \in [0, 1/2) \\ -1 & x \in [1/2, 1) \\ 0 & \text{otherwise} \end{cases}.$$

$S$ is a Hilbert basis; the *Haar system*.

(iii) $L^2(\mathbb{R}, \mu(\mathrm{d}x))$, where $\mu(\mathrm{d}x) = (2\pi)^{-1/2} \exp(x^2/2) \mathrm{d}x$; the Gauss measure. Then take $\{H_n\}_{n \in \mathbb{Z}_{\geq 0}}$, where the $H_n$ are obtained by applying Gram-Schmidt to $\{1, x, x^2, \ldots\}$; the Hermite polynomials. Then $\{H_n\}_{n \in \mathbb{Z}_{\geq 0}}$ is a Hilbert basis.

**Theorem** (Reisz representation theorem)**.** *For any bounded linear functional $\Lambda : H \to \mathbb{R}$ (respectively $\mathbb{C}$), there is a unique $w \in H$ such that $\Lambda(u) = \langle w, u \rangle$ for all $u \in H$.*

## 2.2 Radon-Nikodym Theorems

**Definition.** Let $(E, \mathcal{E})$ be a measurable space and let $\mu, \nu$ be two measures on $(E, \mathcal{E})$. Then $\nu$ is said to be *absolutely continuous* with respect to $\mu$, written $\nu \ll \mu$, if for all $A \in \mathcal{E}$, $\nu(A) = 0$ whenever $\mu(A) = 0$.Two measures $\mu, \nu$ are *mutually singular*, written $\mu \perp \nu$ if there is $B \in \mathcal{E}$ such that $\mu(B) = 0 = \nu(B^c)$.

**Theorem** (Radon-Nikodym)**.** *Let $\mu$ and $\nu$ be finite measures on $(E, \mathcal{E})$ with $\nu \ll \mu$. Then there exists $\omega \in L^1(E, \mathcal{E})$ such that for all $A \in \mathcal{E}$,*

$$\nu(A) = \int_A \omega \mathrm{d}\mu.$$

*Equivalently, for all $h : E \to [0, \infty]$ Borel,*

$$\int h \mathrm{d}\nu = \int h\omega \mathrm{d}\mu.$$

*Proof.* Set $\alpha = \mu + 2\nu$ and $\beta = 2\mu + \nu$. Define

$$\Lambda(f) = \int_E f \mathrm{d}\beta.$$

Then

$$|\Lambda(f)| \leq \int_E |f| \mathrm{d}\beta \leq 2 \int_E |f| \mathrm{d}\alpha \leq 2\sqrt{\alpha(E)} ||f||_{L^2(E,\alpha)}.$$

So $\Lambda : L^2(E, \alpha) \to \mathbb{R}$ is bounded and linear. So by the Riesz representation theorem, there is $g \in L^2(E, \alpha)$ such that $\Lambda(f) = \langle g, f \rangle_{L^2(E,\alpha)}$ for all $f \in L^2(E, \alpha)$. Hence $\int f \mathrm{d}\beta = \int gf \mathrm{d}\alpha$, and

$$\int f(2\mathrm{d}\mu + \mathrm{d}\nu) = \int gf(\mathrm{d}\mu + 2\mathrm{d}\nu) \iff \int f(2 - g)\mathrm{d}\mu = \int f(2g - 1)\mathrm{d}\nu.$$

We claim that $g$ takes values in $[1/2, 2]$ $\mu$-a.e and $\nu$-a.e, and that $g \neq 1/2$ $\mu$-a.e (this implies $g \neq 1/2$ $\nu$-a.e since $\nu \ll \mu$). Assuming the claim, the proof is completed as follows; by the monotone convergence theorem, $(*)$ can be extended to all $f : E \to [0, \infty]$. Given $h : E \to [0, \infty]$ measurable, set

$$f(x) = \frac{h(x)}{2g(x) - 1}, \ \omega(x) = \frac{2 - g(x)}{2g(x) - 1}, \ x \in \{g \neq 1/2\}. \qquad (*)$$

Then

$$\int h \mathrm{d}\nu = \int f(2g - 1)\mathrm{d}\nu = \int f(2g - 1)\mathrm{d}\mu = \int h\omega \mathrm{d}\mu.$$

In particular, taking $h = 1$, we see $\omega \in L^1(E, \mu)$.

Now we prove the claim: let $f = \mathbb{1}_{A_j}$, with $A_j = \left\{x \in E : g(x) < \frac{1}{2} - \frac{1}{j}\right\}$. Then we have

$$\int f(2g - 1)\mathrm{d}\nu \leq -\frac{2}{j}\nu(A_j),$$

$$\int f(2-g)\mathrm{d}\mu \geq \frac{3}{2}\mu(A_j),$$

$$\implies \frac{3}{2}\mu(A_j) \leq -\frac{2}{j}\nu(A_j) \implies \mu(A_j) = \nu(A_j) = 0.$$

Implying $g \geq 1/2$ both $\mu$-a.e and $\nu$-a.e. To show $g \leq 2$ $\mu$-a.e and $\nu$-a.e the proof is analogous, instead with $A_j = \{x \in E : g(x) \geq 2 + 1/j\}$. To show $\mu(\{g = 1/2\}) = 0$, set $f = \mathbb{1}_Z$, $Z = \{g = 1/2\}$ in $(*)$, giving

$$\frac{3}{2}\int \mathbb{1}\{g = 1/2\}\mathrm{d}\mu = 0.$$

$\square$

## 2.3 The dual of $L^p$

**Definition.** A *topological vector space* (TVS) $X$ is a vector space together with a topology in which $(x,y) \mapsto x+y$ and $(\lambda, x) \mapsto \lambda x$ are continuous. The *dual space* $X'$ is the linear space of continuous linear maps $\Lambda : X \to \mathbb{R}$ (or $\mathbb{C}$).

If $X$ is a normed vector space equipped with the topology induced by the norm, then linear maps on $X$ are bounded if and only if they are continuous. We can define a norm on $X'$ by

$$||\Lambda||_{X'} = \sup_{\substack{x \in X \\ ||x|| \leq 1}} |\Lambda(x)|.$$

Then $X'$ is a Banach space (even if $X$ isn't).

We aim to identify $L^p(\mathbb{R}^n)'$ with $L^q(\mathbb{R}^n)$, where $\frac{1}{p} + \frac{1}{q} = 1$, if $p \in [1, \infty)$.

**Proposition.** Let $q \in [1, \infty]$. For every $g \in L^q(\mathbb{R}^n)$,

$$\Lambda_g(f) = \int \bar{f} g \mathrm{d}x$$

defines $\Lambda_g \in L^p(\mathbb{R}^n)'$ with $||\Lambda_g|| = ||g||_{L^q}$.

*Proof.* By Hölder's inequality, $|\Lambda_g(f)| = ||f||_{L^p}||g||_{L^q}$. Hence $\Lambda_g \in L^p(\mathbb{R}^n)'$ and $||\Lambda_g|| \leq ||g||_{L^p}$. Equality: see Example sheet 1. □

**Corollary.** *The map $J : L^q(\mathbb{R}^n) \to L^p(\mathbb{R}^n)'$, $g \mapsto \Lambda_g$ is a linear isometry. Thus we can identify $L^q(\mathbb{R}^n)$ as a subspace of $L^p(\mathbb{R}^n)'$.*

**Remark.** When $p = 2$ then $L^2(\mathbb{R}^n)' = L^2(\mathbb{R}^n)$, i.e $J$ is surjective (Riesz representation theorem).

**Theorem.** *Let $p \in [1, \infty)$, $\frac{1}{p} + \frac{1}{q} = 1$. Then $J$ is surjective, i.e $L^p(\mathbb{R}^n)' = L^q(\mathbb{R}^n)$.*

**Remarks.**

1. $L^1(\mathbb{R}^n)' = L^\infty(\mathbb{R}^n)$, but $L^\infty(\mathbb{R}^n)' \neq L^1(\mathbb{R}^n)$;
2. The same is true if $\mathbb{R}^n$ is replaced by $U \subseteq \mathbb{R}^n$ open.

**Definition.** $\Lambda \in L^p(\mathbb{R}^n)'$ is *positive* if

$$\Lambda(f) \geq 0 \text{ for all } f \in L^p(\mathbb{R}^n) \text{ such that } f \geq 0 \text{ a.e.}$$

**Lemma.** *Let $\Lambda \in L^p(\mathbb{R}^n)'$ be positive. Then there is $g \in L^q(\mathbb{R}^n)$ non-negative with*

$$\Lambda(f) = \int_{\mathbb{R}^n} fg \mathrm{d}x \textit{ for all } f \in L^p(\mathbb{R}^n).$$

*Furthermore $||g||_{L^q} = ||\Lambda||$.*

*Proof.* Let $\mu(\mathrm{d}x) = e^{-|x|^2}\mathrm{d}x$. Then $\mu(\mathbb{R}^n) < \infty$. Define

$$\nu(A) = \Lambda\left(e^{-|x|^2/p}\mathbb{1}_A\right) \text{ for } A \in \mathcal{B}(\mathbb{R}^n).$$

First we show that $\nu$ is a finite measure on $\mathbb{R}^n$. Clearly $\nu(\emptyset) = 0$ and $\nu(A) \in [0, \infty)$ since $\Lambda$ is positive. Let $A_k \in \mathcal{B}(\mathbb{R}^n)$ be a sequence of disjoint sets and $B_m = \bigcup_{k=1}^m A_k$. Then

$$\begin{aligned}|\nu(B_\infty) - \nu(B_m)| &\le \|\Lambda\| \left\| e^{-|x|^2/p}(\mathbb{1}_{B_\infty} - \mathbb{1}_{B_m})\right\|_{L^p} \\ &= \|\Lambda\| \mu(B_\infty \setminus B_m)^{1/p} \to 0.\end{aligned}$$

So $\nu$ is countably additive, and thus a measure. Now we claim $\nu \ll \mu$. Indeed if $\mu(A) = 0$, $\nu(A) \le ||\Lambda||\mu(A)^{1/p}$. Thus by the Radon-Nikodym theorem, there is $\omega \in L^1(\mathbb{R}^n, \mu)$ non-negative such that

$$\nu(A) = \int_A \omega \mathrm{d}\mu = \int_A \omega e^{-|x|^2}\mathrm{d}x \text{ for all } A \in \mathcal{B}(\mathbb{R}^n).$$

Now let $f = e^{-|x|^2/p}\tilde{f}$ where $\tilde{f}$ is simple. Then by linearity of $\Lambda$,

$$\begin{aligned}\Lambda(f) = \int \tilde{f}\mathrm{d}\nu &= \int \tilde{f}\omega e^{-|x|^2}\mathrm{d}x \\ &= \int f \underbrace{\omega e^{-\left(1-\frac{1}{p}\right)|x|^2}}_{\tilde{\omega} = \omega e^{-\frac{1}{q}|x|^2}}\mathrm{d}x.\end{aligned}$$

Hence $\Lambda(f) = \int f\tilde{\omega}\mathrm{d}x$ for all $f$ as above. Exercise: functions of the form $f = e^{-|x|^2/p}\tilde{f}$ for $\tilde{f}$ are dense in $L^p(\mathbb{R}^n)$. Then we have $\Lambda(f) = \int f\tilde{\omega}\mathrm{d}x$ for all $f \in L^p(\mathbb{R}^n)$ since $\Lambda$ is continuous.

Example sheet 1 gives that

$$||\tilde{\omega}||_{L^q} = \sup\left\{\int |f\tilde{\omega}|\mathrm{d}x : ||f||_{L^p} \le 1\right\}.$$

Thus

$$||\tilde{\omega}||_{L^p} \le ||\Lambda|| \text{ since } \int |f\tilde{\omega}|\mathrm{d}x = \int |f|\tilde{\omega}\mathrm{d}x = \Lambda(|f|) \le ||\Lambda||||f||_{L^p}.$$

Convsersely, $\Lambda(f) \le ||f||_{L^p}||\tilde{\omega}||_{L^q}$ by Hölder's inequality, so $||\Lambda|| \le ||\tilde{\omega}||_{L^q}$ and $||\Lambda|| = ||\tilde{\omega}||_{L^q}$. $\square$

**Theorem.** *Let $p \in [1, \infty)$. Then $\int : L^q(\mathbb{R}^n) \to L^p(\mathbb{R}^n)'$, $g \mapsto \Lambda_g$ where $\Lambda_g(f) = \int fg$ is a linear isometry and surjective.*

*Proof.* First consider the real case. In Example Sheet 2 its shown that if $\Lambda \in L^p(\mathbb{R}^n)'$ is real-values, there are $\Lambda_+$ and $\Lambda_-$ both bounded and positive such that $\Lambda = \Lambda_+ - \Lambda_-$. The claim follows from the previous lemma.

In the complex case, if $\Lambda \in L(\mathbb{R}^n, \mathbb{C})'$ then $\Lambda_r(f) = \Re\Lambda(f)$ and $\Lambda_i(f) = \Im\Lambda(f)$ define two $\mathbb{R}$-linear $\Lambda \in L^p(\mathbb{R}^n, \mathbb{R})$ such that

$$\Lambda(f_r + if_i) = \Lambda_r(f_r) - \Lambda_i(f_i) + i\Lambda_r(f_i) + i\Lambda_i(f_r).$$

The claim then follows by the real-valued case. □

## 2.4 Riesz-Markov Theorem

**Fact.** For any finite (positive) regular Borel measure on $\mathbb{R}^n$, $\Lambda_\mu(f) = \int f \mathrm{d}\mu$ defines a positive bounded linear functional on $C_c(\mathbb{R}^n, ||\cdot||_\infty)$.

**Lemma.** *$\Lambda$ uniquely determines $\mu$ and for any $U \in \mathbb{R}^n$ open*

$$\mu(U) = \sup\{\Lambda_\mu(g) : g \in C_c(\mathbb{R}^n), 0 \leq g \leq \mathbb{1}_U\}. \tag{$*$}$$

*Proof sketch.* We would like to take $f = \mathbb{1}_A$ for $A \in \mathcal{B}(\mathbb{R}^n)$, but this is not continuous. So we approximate by continuous functions: assume $U \in \mathbb{R}^n$ is open, set $U_k = U \cap \{|x| < k\}$, abd

$$\chi_k(x) = \begin{cases} 1 & x \in U_k,\ d(x, U_k^c) \geq \frac{1}{k} \\ 0 & x \notin U_k \\ kd(x, U_k^c) & x \in U_k,\ d(x, U_k^c) < 1/k \end{cases}.$$

Then $\chi_k \in C_c(\mathbb{R}^n)$ and $\chi_k \uparrow \mathbb{1}_U$. So by the Monotone Convergence Theorem,

$$\mu(U) = \lim_{k\to\infty} \int \chi_k \mathrm{d}\mu = \lim_{k\to\infty} \Lambda(\chi_k).$$

And $(*)$ also follows. Since $\mu$ is regular, this determines $\mu$ on all Borel sets. □

**Definition.** A *signed measure* is the difference of two mutually singular finite positive measures.

**Theorem** (Riesz-Markov Theorem)**.** *Given $\Lambda : C_c(\mathbb{R}^n) \to \mathbb{R}$ linear positive and bounded, there is a unique finite Borel measure $\mu$ on $\mathbb{R}^n$ such that*

$$\Lambda(f) = \int_{\mathbb{R}} f \mathrm{d}\mu,\ \forall f \in C_c(\mathbb{R}^n).$$

*The dual space $C_c(\mathbb{R}^n)$ is the space of signed measures.*

## 2.5 Strong, weak & weak-$*$ topologies

Example Sheet 2: if $X$ is a Banach space, then the closed unit ball is compact iff $X$ is finite dimensional.

Goal: recover some form of compactness by considering a weaker topology.

**Definition.** A *seminorm* $p$ on a vector space $X$ (over $\mathbb{R}$ or $\mathbb{C}$) is a map $p : X \to \mathbb{R}$ such that

(i) $p(x+y) \leq p(x) + p(y)$ for all $x, y \in X$;

(ii) $p(\lambda x) = |\lambda| p(x)$ for all $x \in X$;

(iii) $p(x) \geq 0$ for all $x \in X$.

(Note: it is not necessarily positive semidefinite)

**Definition.** A family $\mathcal{P}$ of seminorms is *separating* if for every $x \in X$ with $x \neq 0$ there is $p \in \mathcal{P}$ such that $p(x) \neq 0$.

**Definition.** The topology $\tau_{\mathcal{P}}$ induced by a family of seminorms $\mathcal{P}$ is generated by

$$\beta = \{x + B : x \in X,\ B \in \dot{\beta}\}.$$

Where $\dot{\beta}$ consists of finite intersections of $V(p, n) = \{x \in X : p(x) < 1/n\}$ for $p \in \mathcal{P}$, $n \in \mathbb{N}$. $(X, \tau_{\mathcal{P}})$ is a locally convex topological vector space (LCTVS).

**Theorem.** *$\beta$ is a neighbourhood base for the topology $\tau_{\mathcal{P}}$ (every open set $U \in \tau_{\mathcal{P}}$ is a union of sets in $\beta$), and the vector space operations $(x, y) \mapsto x + y$ and $(\lambda, x) \mapsto \lambda x$ are continuous, as is every seminorm $p \in \mathcal{P}$.*

Example Sheet 2: for $(x_k)_{k \geq 1}$ in $X$, $x_k \to x$ in $\tau_{\mathcal{P}}$ if and only if $p(x - x_k) \to 0$ for all $p \in \mathcal{P}$.

**Fact.** If $\mathcal{P} = (p_k)_{k \in \mathbb{N}}$ is countable, then the topology is induced by the metric

$$d_{\mathcal{P}}(x, y) = \sum_{k=1}^{\infty} 2^{-k} \frac{p_k(x-y)}{1 + p_k(x-y)}.$$

**Definition.** If $\mathcal{P}$ is as above with metric as above, if the metric $d_{\mathcal{P}}$ is complete, $(X, d_{\mathcal{P}})$ is called a *Frechet space.*

**Examples.**

(i) $X$ a Banach space, $\mathcal{P}_s = \{||\cdot||\}$: the corresponding topology $\tau_s = \tau_{\mathcal{P}_s}$ is called *norm* or *strong topology*. We have $x_k \to x$ in $\tau_s$ if and only if $||x_k - x|| \to 0$.

(ii) $X$ a Banach space, $\mathcal{P}_w = \{p_\Lambda : \Lambda \in X^1\}$ where $p_\Lambda(x) = |\Lambda(x)|$. Each $p_\Lambda$ is a seminorm and is the Hahn-Banach theorem implies $\mathcal{P}_w$ is separating. (For $X = L^p(\mathbb{R}^n)$ this can be verified directly.) The topology $\tau_w = \tau_{\mathcal{P}_w}$ is called the *weak topology*. We have $x_k \to x$ in $\tau_w$ if and only if $\Lambda(x_k) \to \Lambda(x)$ for all $\Lambda \in X'$. We write $x_k \rightharpoonup^w x$. Also $x_k \to x$ implies $x_k \rightharpoonup^w x$.

(iii) $X$ a Banach space, then $X'$ is also a Banach space. Hence we have a strong and weak topology on $X'$. The *weak-$*$* topology $\tau_{w^*}$ is generated by $\mathcal{P}_{w^*} = \{p_x : x \in X\}$ where $p_x(\Lambda) = |\Lambda(x)|$. Then $\Lambda_k \to \Lambda$ in $\tau_{w^*}$ if and only if $\Lambda_k(x) \to \Lambda(x)$ for every $x \in X$. We write $\Lambda_k \rightharpoonup^{w^*} \Lambda$.

**Remark.** If $X$ is reflexive, i.e $X'' = X$, then $\tau_w = \tau_{w^*}$.

**Example.** Let $p \in [1, \infty)$ and $(f_k)_{k\geq 1}$ be a sequence in $L^p(\mathbb{R}^n)$. Then

$$\begin{aligned} f_k \to f \text{ in } L^p &\iff \int |f_k - f|^p \mathrm{d}x \to 0 \\ f_k \rightharpoonup^w f \text{ in } L^p &\iff \int g(f_k - f)\mathrm{d}x \to 0 \text{ for all } g \in L^q \\ f_k \rightharpoonup^{w^*} f \text{ in } L^p &\iff f_k \rightharpoonup^w f \text{ in } L^p \end{aligned}$$

On the other hand, if $(f_k)_{k\geq 1}$ is in $L^\infty(\mathbb{R}^n)$,

$$\begin{aligned} f_k \to f \text{ in } L^\infty &\iff \text{esssup}|f_k - f| \to 0 \\ f_k \xrightarrow{w^*} \text{ in } L^\infty &\iff \int g(f_k - f)\mathrm{d}x \to 0 \text{ for all } g \in L^1 \\ f_k \xrightarrow{w^*} \text{ in } L^\infty &\iff f_k \xrightarrow{w} \text{ in } L^\infty \end{aligned}$$

## 2.6 Compactness

**Theorem** (Arzela-Ascoli Theorem)**.** *Let $I = [0, 1]$ (or a compact Hausdorff space). Suppose a sequence of continuous functions $f_k : I \to \mathbb{R}$ is*

- *Bounded:* $\sup_k \sup_{x\in I} |f_k(x)| < \infty$
- *Equicontinuous: for all $\varepsilon > 0$ there exists $\delta$ such that* $\sup_k \sup_{x\in I} \sup_{y\in B(x,\varepsilon)} |f_k(x) - f_k(y)| < \varepsilon$.

*Then there is a subsequence $(i_k)$ such that $f_{i_k}$ converges to some continuous $f$.*

Application: $C^{0,\alpha}(I)$ embeds compactly into $C^0(I)$, where $C^{0,\alpha}(I) = \{f \in C^0(I) : ||f||_{C^{0,\alpha}} < \infty\}$,

$$||f||_{C^{0,\alpha}} = \sup_{x \in I} |f'(x)| + \sup_{\substack{x,y \in I \\ x \neq y}} \frac{|f(x) - f(y)|}{|x-y|^\alpha}.$$

The identity map id : $C^{0,\alpha}(I) \to C^0(I)$ is compact, i.e any sequence $(f_i)_{i\geq 1}$ in $C^{0,\alpha}$ that is bounded in $C^{0,\alpha}$ has a convergent subsequence in $C^0(I)$.

**Theorem** (Banach-Alaoglu)**.** *Let $X$ be a separable Banach space, and let $(\Lambda_j)_{j\geq 1}$ be a bounded sequence in $X'$, say $\sup_j ||\Lambda_j||_{X'} \leq 1$. Then there is a subsequence $(j_i)$ and $\Lambda \in X'$ such that $\Lambda_{j_i} \to^{w^*} \Lambda$.*

**Example.** Let $p \in (1, \infty]$ and $(f_j)_{j\geq 1}$ be a sequence in $L^p(\mathbb{R}^n)$ such that $||f_j||_{L^p} \leq K$ for all $j$. Then there is $f \in L^p$ with $||f||_{L^p} \leq K$ and a subsequence $(j_i)$ such that for every $g \in L^q(\mathbb{R}^n)$, $\int f_{j_i} g \mathrm{d}x \to \int fg \mathrm{d}x$. (Just apply Banach-Alaoglu noting $L^q(\mathbb{R}^n)' = L^p(\mathbb{R}^n)$ for $p \in [1, \infty)$ and $L^q$ is separable for such $q$.)

*Proof.* Step 1: construction. Let $D = \{x_k\}_{k=1}^\infty \subseteq X$ be dense (can do this by separability). Since $(\Lambda_j(x_1))_{j\geq 1}$ is a bounded sequence, there is a subsequence $J_1 \subseteq \mathbb{N}$ and $\Lambda(x_1) \in \mathbb{R}$ (or $\mathbb{C}$) such that $\Lambda_j(x_1) \to \Lambda(x_1)$ for $j \in J_1$, $j \to \infty$. Iterating, there are nested subsequences $J_1 \supseteq J_2 \supseteq J_3 \supseteq \ldots$ and $\Lambda(x_k) \in \mathbb{R}$ (or $\mathbb{C}$) such that $\Lambda_j(x_k) \to \Lambda(x_k)$ for $j \in J_l$, $l \geq k$.

Now take the 'diagonal subsequence' $J$ of $J_1 \supseteq J_2 \supseteq \ldots$ defined by $J = (j_n)_{n\geq 1}$ where $j_n$ is the first element of $J_n$. i.e it has first element which is the first element of $J_1$, second element which is the first element of $J_2$, etc. Then $\Lambda_j(x_k) \to \Lambda(x_k)$ for $j \in J, j \to \infty$. $\square$