

1 Lebesgue Integration Theory

1.1 Review of measure theory

Definition. Given a set E , a σ -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 σ -algebra containing \mathcal{A} .

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

- (i) $\mu(\emptyset) = 0$;
- (ii) $A_n \in \mathcal{E}, n \in \mathbb{N} \text{ disjoint} \Rightarrow \mu(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n \in \mathbb{N}} \mu(A_n)$.

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

Definition (Borel measure). If (E, τ) 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. $E = \mathbb{R}^n$, μ the Lebesgue measure satisfying $\mu((a_1, b_1) \times \dots \times (a_n, b_n)) = (b_1 - a_1) \dots (b_n - a_n)$.

Notation: we write $\mu(dx) = dx$ and $\mu(A) = |A|$ when μ is the Lebesgue measure.

Definition (Measurable function). Let (E, \mathcal{E}) and (F, \mathcal{F}) be measurable spaces. Then $f : E \rightarrow F$ 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 $f : E \rightarrow F$ is called a *non-negative measurable function*.

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

Definition. $f : E \rightarrow F$ ($F = [0, \infty]$ 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

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

For a non-negative measurable f , we define

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

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

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

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

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

- (i) $f_n \rightarrow f$ pointwise almost everywhere;
- (ii) $|f_n| \leq g$ almost everywhere for some integrable g .

Then

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

1.2 L^p spaces

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

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

and

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

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

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

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 $E = \mathbb{R}^n$, μ 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 $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$, the *convolution* $f * g$ is defined by

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

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 $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}_{\geq 0}^n$. Set $|\alpha| = \alpha_1 + \dots + \alpha_n$; $\alpha! = \alpha_1! \dots \alpha_n!$; $X^\alpha = X_1^{\alpha_1} \dots X_n^{\alpha_n}$ for $X \in \mathbb{R}^n$; $\nabla^\alpha f = D^\alpha f = \frac{\partial^\alpha f}{\partial x^\alpha} = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}$.

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

Proposition. Let $f \in L_{\text{loc}}^1(\mathbb{R}^n)$, $g \in C_c^k(\mathbb{R}^n)$, 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 $T_z g(x) \rightarrow g(x)$ for all x as $z \rightarrow 0$ (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 $f \in L_{\text{loc}}^1(\mathbb{R}^n)$, $|f(y)|\mathbb{1}_{B_R(0)}(x - y)$ is integrable in y , so by the dominated convergence theorem,

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

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 $\nabla_i^h g(x) \rightarrow \nabla_i g(x)$ as $h \rightarrow 0$.

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, $\nabla_i^h(f * g) = f * (\nabla_i^h g) \rightarrow f * \nabla_i g$. Thus $f * g \in C^1$. The case $k > 1$ is similar, with induction. \square

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

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

Proof. Example sheet 1. \square

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

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

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 $\int_{\mathbb{R}^n} \varphi dx = 1$ 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 $\varphi_\varepsilon * g \rightarrow g$ in L^p .

Proof. We have

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

Hence

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

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

Definition. φ 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 $\|f - f\mathbb{1}_{B_R(0)}\|_{L^p} \rightarrow 0$ as $R \rightarrow \infty$ 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 . \square

1.3 Lebesgue Differentiation Theorem

Recall:

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

We actually have a stronger result:

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

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

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 $G(x) = \int_{-\infty}^x g(t)dt$, then G is differentiable for almost every x with $G'(x) = g(x)$.

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

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

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

Remark. We sometimes write $\int_{B_r(x)} |f(y)|dy$ for $\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)|dy$.

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

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

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

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

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

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

$$\frac{1}{|B_{r_x}(x_k)|} \int_{B_{r_x}(x_k)} |f(y)|dy \rightarrow \frac{1}{|B_{r_x}(x)|} \int_{B_{r_x}(x)} |f(y)|dy.$$

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_λ^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 $\frac{1}{|B_i|} \int_{B_i} |f(y)| dy > \lambda$, it follows that $|B_i| < \frac{1}{\lambda} \int_{B_i} |f(y)| dy$. Hence

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

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

Now we are ready to prove:

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

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

The x for which this holds are called Lebesgue points.

Proof. Let

$$A_\lambda = \left\{ x \in \mathbb{R}^n : \limsup_{r \rightarrow 0} \int_{B_r(x)} |f(y) - f(x)| dy > 2\lambda \right\}$$

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

$$\begin{aligned} \int_{B_r(x)} |f(y) - f(x)| dy &\leq \underbrace{\int_{B_r(x)} |f(y) - g(y)| dy}_{\leq M(f-g)(x)} + |f(x) - g(x)| + \underbrace{\int_{B_r(x)} |g(x) - g(y)| dy}_{\rightarrow 0 \text{ since } g \in C^\infty} \\ \implies \limsup_{r \rightarrow 0} \int_{B_r(x)} |f(y) - f(x)| dy &\leq M(f-g)(x) + |f(x) - g(x)|. \end{aligned}$$

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

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

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

1.4 Littlewood's Principles

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

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

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

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

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

Now it suffices to show $f_j \rightarrow f$ 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 $\limsup_{j \rightarrow \infty} \sup_{A_N} |f_j - f| \leq \frac{1}{n}$ for all $n \geq N$, hence $\lim_{j \rightarrow \infty} \sup_{A_N} |f_j - f| = 0$. \square

Theorem (Lusin). *Let $f : E \rightarrow \mathbb{R}$ 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 not 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 $f_n \rightarrow f$ 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_ε such that $f_n \rightarrow f$ uniformly on A_ε 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 $f_n \rightarrow f$ uniformly on F'_ε , $f|_{F'_\varepsilon}$ 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. \square

2 Banach and Hilbert space analysis

2.1 The Hilbert space L^2

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

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

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 $\overline{\text{span}\{u_j\}} = H$.

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^∞).
- (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) $L^2(\mathbb{R}, \mu(dx))$, where $\mu(dx) = (2\pi)^{-1/2} \exp(x^2/2) dx$; the Gauss measure. Then take $\{H_n\}_{n \in \mathbb{Z}_{\geq 0}}$, where the H_n are obtained by applying Gram-Schmidt to $\{1, x, x^2, \dots\}$; 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 $\Lambda : H \rightarrow \mathbb{R}$ (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 μ, ν be two measures on (E, \mathcal{E}) . Then ν is said to be *absolutely continuous* with respect to μ , written $\nu \ll \mu$, if for all $A \in \mathcal{E}$, $\nu(A) = 0$ whenever $\mu(A) = 0$. Two measures μ, ν 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 μ and ν 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 $A \in \mathcal{E}$,*

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

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

$$\int h d\nu = \int h \omega d\mu.$$

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

$$\Lambda(f) = \int_E f d\beta.$$

Then

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

So $\Lambda : L^2(E, \alpha) \rightarrow \mathbb{R}$ 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 $\int f d\beta = \int g f d\alpha$, and

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

We claim that g takes values in $[1/2, 2]$ μ -a.e and ν -a.e, and that $g \neq 1/2$ μ -a.e (this implies $g \neq 1/2$ ν -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 $f : E \rightarrow [0, \infty]$. Given $h : E \rightarrow [0, \infty]$ measurable, set

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

Then

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

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 $A_j = \{x \in E : g(x) < \frac{1}{2} - \frac{1}{j}\}$. Then we have

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

$$\int f(2 - g) d\mu \geq \frac{3}{2} \mu(A_j),$$
$$\implies \frac{3}{2} \mu(A_j) \leq -\frac{2}{j} \nu(A_j) \implies \mu(A_j) = \nu(A_j) = 0.$$

Implying $g \geq 1/2$ both μ -a.e and ν -a.e. To show $g \leq 2$ μ -a.e and ν -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

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

□