Lecture Notes in Measure and Integration

Lectures by Dr. Alexander Panchenko

Throughout this document, □ signifies end proof, and ▲ signifies end of example.

Table of Contents

Notes by Jakob Streipel. Last updated December 8, 2017.

Lecture 1 Laying the Ground Works

1.1 σ-Algebras

A measure μ takes a set A and maps it onto a nonnegative real number $\mu(A)$. Through trial and error over the years it was found that it is not helpful to have a measure on all subsets of a space; instead we need a certain structure.

Definition 1.1.1 (σ-algebra)**.** A ***σ-algebra*** Σ is a collection of subsets of a fixed set (or universe) Ω such that

(*i*) $\Omega \in A$;

(*ii*) if $A \in \Sigma$, then $A^c \in \Sigma$ (here complement is relative to the universe, i.e. $A^c = \Omega - A$);

(*iii*) if $A_1, A_2, \ldots, A_n$ are in Σ, then $\bigcup_{i=1}^{\infty} A_i \in \Sigma$.

Note that (*i*) and (*ii*) imply that $\varnothing \in \Sigma$, and by De Morgan (*ii*) and (*iii*) imply that Σ is closed under countable intersections as well.

It requires slightly more work, but it is also true that Σ is closed under set difference.

Usually in practice we generate σ-algebras from a (relatively) small family of subsets $\mathcal{F}$.

Example 1.1.2. Consider $\mathcal{F}$ as the family of open subsets of $\mathbb{R}^n = \Omega$. Then define $\mathcal{B}$ to be the smallest σ-algebra containing $\mathcal{F}$. Then $\mathcal{B}$ is called the ***Borel σ-algebra***. Clearly this construction is possible not only for $\mathbb{R}^n$, but for any topological space, giving us a sense of open sets. ▲

Definition 1.1.3 (Measure)**.** A measure μ is a function from a σ-algebra Σ to the nonnegative reals (possibly with ∞), such that

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

(*ii*) $\mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mu(A_i)$ if $A_1, A_2, \ldots, A_n \in \Sigma$ and A_i are disjoint.

Examples 1.1.4. One simple measure is the Dirac measure on $\Omega = \mathbb{R}^n$. This is defined as δy for $y \in \mathbb{R}^n$ where y is a fixed point, and

$$\delta y(A) = \begin{cases} 1, & \text{if } y \in A \\ 0, & \text{otherwise} \end{cases}$$

for $A \in \Sigma$. This is sometimes also called the counting measure, especially if generalised to a greater range and more points. For this purpose it doesn't matter very much what the σ-algebra is, since it cares only about the point y.

Another typical measure is the Lebesgue measure on $\Omega = \mathbb{R}^n$, wherein $\Sigma = \mathcal{B}$, the Borel σ-algebra. Here $\mu := \mathcal{L}^n$ coincides with Euclidean volume for balls and rectangles.

(Recall, for fun, that

$$\mathcal{L}^n(B_{x,r}) = \mathcal{L}^n(B_{0,1})r^n = \frac{2\pi^{n/2}}{n\Gamma(n/2)}r^n,$$

where by $B_{x,r}$ we mean the ball of radius r centred on x.)

Another measure with which most are familiar but might not recognise as such comes from probability theory. Consider some fixed $y_1, y_2, \ldots, i_k \in \mathbb{R}^n$ and some $0 \leq p_1, p_2, \ldots, p_k \leq 1$ such that $\sum_{i=1}^k p_i = 1$. Then

$$\mu = \sum_{i=1}^{k} p_i \delta y_i$$

is a measure with the property that $\mu(\mathbb{R}^n) = 1$. ▲

We end this first lecture with two useful properties of the Lebesgue measure:

(i) Outer regularity, $\mathcal{L}^n(A) = \inf\{\, \mathcal{L}^n(O) \mid A \subset O, O \text{ open} \,\}$; and

(ii) Inner regularity, $\mathcal{L}^n(A) = \sup\{\, \mathcal{L}^n(C) \mid C \subset A, C \text{ compact} \,\}$.

Lecture 2 Measures and Measure Spaces

2.1 Regularity and Other Things

Definition 2.1.1 (Regular measure)**.** Let μ be a measure on some space X. We say that μ is ***regular*** if for every subset $A \subset X$ there exists a measurable set B such that $A \subset B$ and $\mu(A) = \mu(B)$.

Definition 2.1.2 (Borel and Radon measure)**.** A measure μ is said to be ***Borel*** if every Borel set is μ-measurable.

Moreover μ is said to be ***Radon*** if

(i) μ is Borel regular, and

(ii) $\mu(K) < \infty$ for every compact $K \subset X$.

Example 2.1.3. An example of a Radon measure is the probability measure, where the measure of the whole space is 1 and the measure of subsets is bounded by 1, and so certainly measures of compact sets are finite. ▲

Definition 2.1.4 (σ-finiteness)**.** A measure space (Ω, Σ, μ) is ***σ-finite*** if there exists a sequence $\{\, A_i \,\}_{i=1}^{\infty}$, $A_i \in \Sigma$, such that $\mu(A_i) < \infty$ for every $i = 1, 2, 3, \ldots$ and $\Omega = \bigcup\limits_{i=1}^{\infty} A_i$.

Example 2.1.5. The Lebesgue measure is σ-finite since we can decompose $\mathbb{R}^n$ into countably many cubes, all with finite volume (= measure). ▲

Example 2.1.6. The counting measure with countably infinite number of charged points in some cube is a measure that is not Radon. ▲

2.2 Some Properties of Measures

Proposition 2.2.1. *Let μ be a measure. Then*

(i) *if $A \subset B$, then $\mu(A) \leq \mu(B)$;*

(ii) if $A_1 \subset A_2 \subset \ldots$, *then* $\lim\limits_{n=\infty} \mu(A_j) = \mu\Big(\bigcup\limits_{i=1}^{\infty} A_i\Big)$;

(iii) if $A_1 \supset A_2 \supset \ldots$ *and* $\mu(A_1) \leq \infty$, *then* $\lim\limits_{j \to \infty} \mu(A_j) = \mu\Big(\bigcap\limits_{i=1}^{\infty} A_i\Big)$.

Proof. (i) Note first of all that $B = (B - A) \cup A$, where those two parts are disjoint. Now since they are disjoint we can use additivity of measures, whereby

$$\mu(B) = \mu((B - A) \cup A) = \mu(B - A) + \mu(A),$$

and since measures are nonnegative this is certainly greater than or equal to $\mu(A)$ alone.

(ii) We use effectively the same trick, except on more than one set. First, though, fixing $i = k$ we have

$$\bigcup_{i=1}^{k-1} A_i \subset A_k \subset \bigcup_{i=1}^{\infty} A_i$$

and so by (i) we have

$$\mu\Big(\bigcup_{i=1}^{k-1} A_i\Big) \leq \mu(A_k) \leq \mu\Big(\bigcup_{i=1}^{\infty} A_i\Big).$$

Now

$$\bigcup_{i=1}^{\infty} A_i = A_1 \cup \Big(\bigcup_{i=2}^{\infty} (A_i - A_{i-1})\Big)$$

and therefore

$$\mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \mu(A_1) + \sum_{i=2}^{\infty} \mu(A_i - A_{i-1})$$

as well as

$$\mu\Big(\bigcup_{i=1}^{k} A_i\Big) = \mu(A_1) + \sum_{i=2}^{k} \mu(A_i - A_{i-1})$$

and the latter is the partial sum of the latter, so they either converge to the same thing or, by being increasing, both diverge to infinity.

(iii) Finally by De Morgan this last point follows from (ii). □

2.3 Product Measure

Given two spaces Ω_1 and Ω_2, with σ-algebras Σ_1 and Σ_2 along with measures μ_1 and μ_2 respectively, we define $\Omega = \Omega_1 \times \Omega_2$.

Now how ought we combine the σ-algebras? Just taking Cartesian product we might not get another σ-algebra, so what we do is define Σ to be the smallest σ-algebra containing each element from the Cartesian product $\Sigma_1 \times \Sigma_2$.

It is a fact that there exists a unique measure μ on Ω such that

$$\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$$

for all $A_1 \in \Sigma_1$ and $A_2 \in \Sigma_2$.

2.4 Monotone Classes

Definition 2.4.1 (Monotone class)**.** A ***monotone class*** $\mathcal{M}$ is a collection of sets such that

(*i*) $A_i \in \mathcal{M}$ for $i = 1, 2, 3, \dots$ and $A_1 \subset A_2 \subset \dots$ implies that $\bigcup_{i=1}^{\infty} A_i \in \mathcal{M}$;

(*ii*) $B_i \in \mathcal{M}$ for $i = 1, 2, 3, \dots$ and $B_1 \supset B_2 \supset \dots$ implies that $\bigcap_{i=1}^{\infty} B_i \in \mathcal{M}$.

Note that this means that a monotone class is almost a σ-algebra, with the caveat that the closure conditions apply only on monotone families of sets.

Definition 2.4.2 (Algebra)**.** A collection of sets $\mathcal{A}$ is an ***algebra*** if for every $A, B \in \mathcal{A}$ we have $A - B, B - A, B \cup A \in \mathcal{A}$.

Theorem 2.4.3 (Monotone class theorem)**.** *Let Ω be a set, $\mathcal{A}$ an algebra of subsets of Ω such that $\varnothing \in \mathcal{A}$ and $\Omega \in \mathcal{A}$. Then there exists a smallest monotone class $\mathcal{S}$ containing $\mathcal{A}$, and in particular this monotone class is also a σ-algebra.*

2.5 Measurable Functions and Integrals

Definition 2.5.1 (Measurable function)**.** Let $f\colon \Omega \to \mathbb{R}$ and Σ a σ-algebra on Ω. Then f is said to be ***measurable*** with respect to Σ if for all $t \in \mathbb{R}$ the set

$$S_f(t) = \{\, x \in \Omega \mid f(x) > t \,\}$$

is measurable, i.e. $S_f(t) \in \Sigma$.

Remark 2.5.2. The $>$ can be replaced by $\geq$, $\leq$, or $<$, for example by taking

$$\{\, x \in \Omega \mid f(x) > t \,\} = \bigcup_{i=1}^{\infty} \{\, x \in \Omega \mid f(x) \geq t + 1/j \,\}.$$

Example 2.5.3. Take $\Sigma = \mathcal{B}$, $\Omega = \mathbb{R}^n$, and f continuous on Ω. Then f is (Borel) measurable. ▲

Lecture 3 Measurable Functions

3.1 Measurability of Functions

We claim that if $f\colon \Omega \to \mathbb{R}$ is continuous with Ω being some open and metrisable space, then

$$S_f(t) = \{\, x \in \Omega \mid f(x) > t \,\}$$

is open for every $t \in \mathbb{R}$.

Proof. Since f is continuous, for every $\varepsilon > 0$ there exists some $\delta > 0$ such that $0 < |x - y| < \delta$ implies $|f(x) - f(y)| < \varepsilon$.

Let $x \in S_f(t)$ and, say, $d = f(x) - t > 0$. Pick $\varepsilon = d/2$, implying that there exists some $\delta_* > 0$ such that $0 < |x - y| < \delta_*$ implies $|f(x) - f(y)| < d/2$.

Then

$$\begin{aligned} f(y) &= f(x) + f(y) - f(y) \geq f(x) - |f(y) - f(x)| > f(x) - \frac{d}{2} \\ &= f(x) - \frac{f(x) - t}{2} = \frac{f(x)}{2} + \frac{t}{2} > t. \end{aligned}$$ □

The same can be shown for semi-continuous functions (i.e. functions for which superlevel sets or sublevel sets are open, but not necessarily both).

All of these are measurable.

Remark 3.1.1. By the above, the set $\{ x \in \Omega \mid f(x) < t + 1/j \}$ is measurable for all $j = 1, 2, \ldots$, and

$$\{ x \in \Omega \mid f(x) \leq t \} = \bigcap_{j=1}^{\infty} \{ x \in \Omega \mid f(x) < t + 1/j \},$$

and since this is a countable intersection of measurable sets, it too is measurable.

Theorem 3.1.2. *Let f and g be measurable functions. Then*

(*i*) $\lambda f + \gamma g$ *is measurable for all* $\lambda, \gamma \in \mathbb{R}$*;*

(*ii*) fg *is measurable;*

(*iii*) $|f|$ *is measurable;*

(*iv*) $\varphi(f(x))$ *is measurable for all Borel measurable* φ *in* $\mathbb{R}$*;*

(*v*) $\max\{ f(x), g(x) \}$ *and* $\min\{ f(x), g(x) \}$ *are measurable;*

(*vi*) $\limsup\{ f_j \}$ *and* $\liminf\{ f_j \}$ *are measurable when* f_j *are measurable.*

Proof sketch. (*i*) Consider

$$\begin{aligned} \{ x \in \Omega \mid f(x) + g(x) > t \} &= \bigcup_{q \in \mathbb{Q}, q < t} \{ x \in \Omega \mid f(x) + g(x) \geq q \} \\ &= \bigcup_{q \in \mathbb{Q}, q < t} \bigcup_{r \in \mathbb{Q}, r \leq q} \{ x \in \Omega \mid f(x) \geq r \} \cup \{ x \in \Omega \mid g(x) \geq q - r \} \end{aligned}$$

where both of the latter are measurable, and we have a countable union of countable unions of measurable sets, so it again is measurable.

The other proofs are similar but arguably easier. We use clever ideas like $fg = ((f+g)^2 - (f-g)^2)/4$ and write the parts in terms of each other. □

Definition 3.1.3 (Strictly positive function)**.** A non-negative measurable function is called ***strictly positive*** if $\{ x \in \Omega \mid f(x) = 0 \} = \ker f$ has measure 0.

Definition 3.1.4 (Support)**.** Let μ be a Borel measure on $\mathbb{R}^n$. Let f be a μ-measurable function. Consider open sets $\omega \in \mathcal{B}$ such that $f(x) = 0$ for almost all $x \in \omega$ (meaning that it is zero everywhere except possibly on subsets of measure 0). Let $\omega^* = \bigcup \omega$, and define $\operatorname{ess\,supp} f = \Omega - \omega^*$. This closed set is called the ***(essential) support*** of f.

With all of this in hand, we can define a meaningful integration in terms of measures.

Definition 3.1.5. Let $f\colon \Omega \to \mathbb{R}^+$, with f being μ-measurable. Define $F_f(t) = \mu(S_f(t))$, which makes F_f nonincreasing. Then we define

$$\int_\Omega f(x)\mu(dx) := \int_0^\infty F_f(t)\,dt.$$

Clearly this might not be finite, but it either converges to a limit or to ∞, by the monotonicity.

Definition 3.1.6 (Summable or Integrable functions)**.** For functions $f \geq 0$ we say that they are μ-***summable*** (or μ-***integrable***) if

$$\int_0^\infty F_f(t)\,dt < \infty.$$

For μ-measurable functions f which aren't nonnegative, we decompose $f(x) = f_+(x) - f_-(x)$, where $f_+(x) = \max\{\, f(x), 0\,\}$ and $f_-(x) = -\min\{\, 0, f(x)\,\}$, and given this f is μ-summable if

$$\int_\Omega f_\pm(x)\mu(dx) < \infty.$$

Lecture 4 Integration

4.1 Properties of Integration

By similar computations as last time for measurability, if f and g are summable, so is $\lambda f + \gamma g$ for all constants λ and γ. Moreover

$$\int_\Omega \lambda f + \gamma g\,\mu(dx) = \lambda \int_\Omega f\,\mu(dx) + \gamma \int_\Omega g\,\mu(dx).$$

A question which often arises in practice is whether we can exchange the order of integration and limits, and when. In other words, what is the interplay of summability and convergence?

Example 4.1.1. Consider a sequence of functions $\{\, f_j\,\}_{j=1}^\infty$, where $f_j \to f$ pointwise, and f_j and f are all Riemann integrable. Then to have

$$\lim_{j\to\infty} \int_a^b f_j(x)\,dx = \int_a^b \lim_{j\to\infty} f_j(x)\,dx = \int_a^b f(x)\,dx$$

a sufficient condition is for f_j to converge to f not only pointwise, but uniformly. ▲

We can get much better results, however, where we don't have to strengthen the convergence that much.

Theorem 4.1.2 (Monotone convergence)**.** *Let $\{\, f_j\,\}_{j=1}^\infty$ be increasing pointwise, with f_j summable on (Ω, Σ, μ). Then*

$$f = \lim_{j\to\infty} f_j$$

pointwise almost everywhere with respect to μ, *and*

$$I = \lim_{j\to\infty} I_j = \lim_{j\to\infty} \int_\Omega f_j \, \mu(dx)$$

is finite if and only if f *is summable, in which case*

$$I = \int_\Omega f \, \mu(fx).$$

In other words we get

$$\lim_{j\to\infty} \int_\Omega f_j \, \mu(dx) = \int_\Omega \lim_{j\to\infty} f_j \, \mu(dx).$$

Proof. Without loss of generality, assume $f_j \geq 0$ (otherwise we work with $f_j - f_1$, and by linearity of integration we can extract the same property afterward). Consider $F_{f_j}(t) = \mu(\{ x \in \Omega \mid f_j(x) > t \})$. Let

$$S = \{ x \in \Omega \mid f(x) > t \} = \bigcup_{j=1}^{\infty} S_j$$

where $S_j = \{ x \in \Omega \,|\, f_j(t) > t \}$. Note that S_j are nested by f_j being increasing, so $S_1 \subset S_2 \subset \ldots$, whereby

$$\lim_{h\to\infty} \mu(S_j) = \mu\Big(\bigcup_{j=1}^{\infty} S_j \Big) = \mu(S)$$

and therefore

$$\lim_{j\to\infty} F_{f_j}(t) = F_f(t)$$

where the sequence on the left is increasing. Thus we ask ourselves whether

$$\lim_{j\to\infty} \int_0^\infty F_{f_j}(t) \, dt = \int_0^\infty F_f(t) \, dt,$$

which by the monotonicity above makes them equal if the right-hand side is bounded, which it is by assumption of f being summable. This completes the forward direction. □

Lemma 4.1.3 (Fatou). *Let* $\{ f_j \}_{j=1}^{\infty}$, *with* $f_j \geq 0$ *and* f_j *summable on* (Ω, Σ, μ). *Then*

$$f(x) = \liminf_{j\to\infty} f_j(x)$$

is measurable and

$$\liminf_{j\to\infty} \int_\Omega f_j(x) \, \mu(dx) \geq \int_\Omega f(x) \, \mu(x)$$

Remark 4.1.4. Note that if $f_j \to f$ almost everywhere, then

$$\liminf_{j\to\infty} \int_\Omega f_j(x) \, \mu(dx) \geq \int_\Omega \lim_{j\to\infty} f_j(x) \, \mu(dx).$$

Example 4.1.5. Consider the function

$$f_j(x) = \begin{cases} 1/j, & |x| \leq j \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$\lim_{j\to\infty} \int_{\mathbb{R}} f_j(x)\,dx = \lim_{j\to\infty} 2 = 2$$

but $f_j \to 0$ pointwise, so

$$\int_{\mathbb{R}} \lim_{j\to\infty} f_j(x)\,dx = \int_{\mathbb{R}} 0\,dx = 0,$$

so clearly the former exceeds the latter. ▲

Proof. Consider

$$F_k(x) = \int_{j\geq k} f_j(x).$$

Then $\{ F_k \}_{k=1}^{\infty}$ is an increasing sequence, and F_k is measurable since

$$\{ x \in \Omega \mid F_k(x) \geq t \} = \bigcap_{j\leq k} \{ x \in \Omega \mid f_j(x) \geq t \}$$

is a countable intersection of measurable sets. Moreover $F_k(x) \leq f_k(x)$ and so since we have nonnegativity we have that $F_k(x)$ are summable. Now

$$\lim_{k\to\infty} F_k(x) = \sup_{k\geq 1} F_k(x) = \sup_{k\geq 1} \inf_{j\geq k} f_j(x) = \liminf_{j\to\infty} f_j(x).$$

Finally

$$\begin{aligned} \liminf_{j\to\infty} \int_\Omega f_j\,\mu(dx) &:= \sup_{k\geq 1} \inf_{j\geq k} \int_\Omega f_j\,\mu(dx) \geq \lim_{k\to\infty} \int_\Omega F_k(x)\,\mu(dx) \\ &= \int_\Omega f(x)\,\mu(dx) = \int_\Omega \liminf_{j\to\infty} f_j\,\mu(dx), \end{aligned}$$

where we use the Monotone convergence theorem on the last line.

Note that the inequality above is using the elementary observation that

$$\inf_j \int h_j \geq \inf_j \int \inf_j h_j = \int \inf_j h_j.$$ □

The tour de force in this type of problem is the Dominated convergence theorem. We will prove this later, but for now we describe the big picture. Suppose $f_j \to f$ almost everywhere. Then if there exists some function G such that $|f_j(x)| \leq G(x)$ almost everywhere, then if G is summable, then the limit of f_j is summable as well, and the integrals of f_j converge to the integral of f.

Theorem 4.1.6 (Dominated convergence)**.** *Let* $\{ f_j \}_{j=1}^{\infty}$ *be a sequence of functions on* (Ω, Σ, μ)*. Then if*

(*i*) f_j *are summable,*

(*ii*) *$f_j \to f$ almost everywhere, and*

(*iii*) *there exists a summable $G(x)$ such that $|f_j(x)| \leq G(x)$ for $j = 1, 2, \ldots$,*

then $|f(x)| \leq G(x)$ and therefore is summable, and

$$\lim_{j\to\infty} \int_\Omega f_j(x)\,\mu(dx) = \int_\Omega f(x)\,\mu(dx).$$

The idea of the proof is to use Fatou's lemma on the functions f_j and $G - f_j$.

Lecture 5 Dominating Convergence

5.1 Proof of Dominating Convergence Theorem

Theorem 5.1.1 (Dominated Convergence)**.** *Let $\{ f_j \}$ be summable functions on (Ω, Σ, μ), and suppose $f_j \to f$ almost everywhere pointwise. Suppose further that there exists some function G on (Ω, Σ, μ) such that $|f_j(x)| \leq G(x)$ for every j. Then $|f(x)| \leq G(x)$ and*

$$\lim_{j\to\infty} \int_\Omega f_j(x)\,\mu(dx) = \int_\Omega f(x)\,\mu(dx).$$

Proof. Since the positive and negative parts of f_j satisfy the same assumptions it suffices to prove the theorem for $f_j \geq 0$.

We get an immediate upper bound by Fatou's lemma;

$$\liminf_{j\to\infty} \int_\Omega f_j(x)\,\mu(dx) \geq \int_\Omega f(x)\,\mu(dx).$$

For a lower bound, consider the function $(G - f_j)(x) \geq 0$. By Fatou's lemma, we have

$$\liminf_{j\to\infty} \int_\Omega (G - f_j)(x)\,\mu(dx) \geq \int_\Omega (G - f)(x)\,\mu(dx)$$

which implies

$$\liminf_{j\to\infty} \int_\Omega -f_j(x)\,\mu(dx) \geq -\int_\Omega f(x)\,\mu(dx),$$

and if we multiply by -1 on either side we have

$$-\liminf_{j\to\infty} -\int_\Omega f_j(x)\,\mu(dx) \leq \int_\Omega f(x)\,\mu(dx).$$

Now the negative $\liminf$ of the negative of a sequence is the $\limsup$ of the opposite sequence, so

$$\limsup_{j\to\infty} \int_\Omega f_j(x)\,\mu(dx) \leq \int_\Omega f(x)\,\mu(dx).$$

Together with the upper bound we have

$$\limsup_{j\to\infty} \int_\Omega f_j(x)\,\mu(dx) \leq \int_\Omega f(x)\,\mu(dx) \leq \liminf_{j\to\infty} \int_\Omega f_j(x),$$

but by definition of $\limsup$ and $\liminf$ we have $\liminf \leq \limsup$, so by anti-symmetry we have

$$\lim_{j\to\infty} \int_\Omega f_j(x)\,\mu(dx) = \int_\Omega f(x)\,\mu(dx).$$ □

Theorem 5.1.2. *Suppose that* $\{ f_j \}_{j=1}^{\infty}$ *is a sequence on functions on* (Ω, Σ, μ) *and that* $f_j \to f$ *almost everywhere pointwise. Suppose further that there exists* $p \in (0, \infty)$ *and* $C > 0$ *such that*

$$\int_{\Omega} |f_j(x)|^p \, \mu(dx) < C$$

for $j = 1, 2, 3, \ldots$. *Then*

$$\lim_{j\to\infty} \int_{\Omega} \Big| |f_j(x)|^p - |f_j(x) - f(x)|^p - |f(x)|^p \Big| \, \mu(dx) = 0.$$

Proof. We will use the following variation of the triangle inequality in the proof, which we will prove later: for every $\varepsilon > 0$ there exists some $C_\varepsilon > 0$ such that

$$||a + b|^p - |b|^p| \le \varepsilon |b|^p + C_\varepsilon |a|^p \tag{5.1.1}$$

for every $a, b \in \mathbb{R}$.

Now with that in hand, we write $f_j = f + g_j$, with $g_j \to 0$ almost everywhere. Consider the function

$$G_{\varepsilon,j} = (|f + g_j|^p - |g_j|^p - |f|^p - \varepsilon |g_j|^p)_+$$

(which we note differs from the integrand in the theorem only in the ε term and the positive part).

Now we claim that the limit of the integral of $G_{\varepsilon,j}$ is 0. To see this, consider

$$||f + g_j|^p - |g_j|^p - |f|^p| \le ||f + g_j|^p - |g_j|^p| + |f|^p \le \varepsilon |g_j|^p + (1 + C_\varepsilon)|f|^p$$

where we use (5.1.1) in the last step. Now this means that $G_{\varepsilon,j} \le (1 + C_\varepsilon)|f|^p$. To see that this converges to 0 now, note that $|f + g_j|^p = |f_j|^p \to |f|^p$ almost everywhere since absolute value and raising to the power p are continuous functions.

Now this means that we may use the dominated convergence theorem to show that the integral of $G_{\varepsilon,j}$ goes to 0.

Moving on, note that

$$\int_{\Omega} |f + g_j|^p - |g_j|^p - |f|^p + \varepsilon |g_j|^p - \varepsilon |g_j|^p \, \mu(dx) \le \varepsilon \int_{\Omega} |g_j|^p \, \mu(dx) + \int_{\Omega} G_{\varepsilon,j} \, \mu(dx)$$

and

$$\int_{\Omega} |g_j|^p \, \mu(dx) = \int_{\Omega} |f - f_j|^p \, \mu(dx) \le 2^p \int_{\Omega} |f|^p + |f_j|^p \, \mu(dx) \le 2^{p+1} C.$$

The reason for 2^{p+1} is because we need it to be true for all positive p. For $p \ge 1$ we have convexity, so 2^{p-1} would suffice, but for $0 < p < 1$ it wouldn't.

More later... □

Lecture 6 Product Measure

6.1 Finishing the Proof

Proof continued. We proved already that

$$\int_{\Omega} ||f + g_i|^p - |g_i|^p - |f|^p| \, \mu(dx) \le \varepsilon \int_{\Omega} |g_i|^p \, \mu(dx) + \int_{\Omega} G_{\varepsilon,i} \, \mu(dx)$$

and also

$$\int_\Omega |g_i|^p \, \mu(dx) \leq 2^{p+1}C$$

and

$$\int_\Omega G_{\varepsilon,i} \, \mu(dx) \to 0$$

by the Dominated convergence theorem.

This implies that

$$\limsup_{j\to\infty} \int_\Omega ||f+g_j|^p - |g_j|^p - |f|^p| \, \mu(dx) \leq 2^{p+1}C\varepsilon + \varepsilon = \varepsilon D.$$

Since $\varepsilon > 0$ is arbitrary, this implies that

$$\limsup_{j\to\infty} \int_\Omega ||f+g_j|^p - |g_j|^p - |f|^p| \, \mu(dx) = 0$$

which in turn implies that

$$\lim_{j\to\infty} \int_\Omega ||f+g_j|^p - |g_j|^p - |f|^p| \, \mu(dx) = 0$$

since we have that the lim sup of a nonnegative quantity is 0.

Now there are some parts remaining: we have yet to show the curious inequality we started with, i.e. that for all $\varepsilon > 0$ there exists some $C_\varepsilon > 0$ such that

$$||a+b|^p - |b|^p| \leq \varepsilon|b|^p - C_\varepsilon|a|^p$$

for all $a, b \in \mathbb{R}$.

Now we are considering a map $t \mapsto |t|^p$, which for $p \geq 1$ is convex, so this case we deal with that way. Write

$$a+b = (1-\lambda)\underbrace{\frac{a}{1-\lambda}}_{=A} + \lambda\underbrace{\frac{b}{\lambda}}_{=B}$$

and so

$$|a+b|^p = |(1-\lambda)A + \lambda B|^p \leq (1-\lambda)|A|^p + \lambda|B|^p = (1-\lambda)^{1-p}|a|^p + \lambda^{1-p}|b|^p.$$

We want $\lambda^{1-p} = \varepsilon + 1$, whereby $(1-\lambda)^{1-p} = C_\varepsilon$, meaning that

$$|a+b|^p \leq (1+\varepsilon)|b|^p + C_\varepsilon|a|^p$$

which, rearranged, is

$$|a+b|^p - |b|^p \leq \varepsilon|b|^p + C_\varepsilon|a|^p.$$

It remains to show that $|b|^p - |a+b|^p$ has the same bound.

If on the other hand we have $0 < p \leq 1$, then we have that $|a+b|^p \leq |a|^p + |b|^p$, where without loss of generality we can assume that $a, b \neq 0$ (since otherwise it's trivial) and the inequality is equivalent with

$$|b|^p|a/b+1|^p \leq |(|a|/|b|)^2 + 1||b|^p$$

where if we let $x = a/b$ this is equivalent to

$$|x+1|^p \leq |x|^p + 1.$$

When $x \geq 0$ it is easy to see that this is true by elementary calculus, and the other case follows similarly. □

6.2 Product Measure

Theorem 6.2.1. *Let $(\Omega_i, \Sigma_i, \mu_i)$ for $i = 1, 2$ be σ-finite measure spaces. Let $A \subset \Sigma_1 \times \Sigma_2$, then A is measurable.*

For all $x \in \Omega_2$ define $f(x) = \mu_1(A_1(x))$ where

$$A_1(x) = \{\, y \in \Omega_1 \mid (x, y) \in A \,\}$$

and for all $y \in \Omega_1$ define $g(y) = \mu_2(A_2(y))$ where

$$A_2(y) = \{\, x \in \Omega_2 \mid (x, y) \in A \,\}.$$

Then f is Σ_2 measurable, g is Σ_1-measurable, and the integrals

$$I = \int_{\Omega_2} f(x)\, \mu_2(dx) = \int_{\Omega_1} g(y)\, \mu_1(dy)$$

define the product measure $\mu_1 \times \mu_2$ by

$$(\mu_1 \times \mu_2)(A) = I.$$

Finally $\mu_1 \times \mu_2$ is σ-finite.

Proof. Consider $\{\, A_i \,\}_{i=1}^{\infty}$ a disjoint set in $\Sigma_1 \times \Sigma_2$.

It is proven in the book that ***sections*** $A_{i,1}(x) = \{\, y \in \Omega_1 \mid (x, y) \in A_i \,\}$ are measurable and disjoint, implying that

$$\mu_1\Big(\bigcup_{i=1}^{\infty} A_{i,1}(x)\Big) = \sum_{i=1}^{\infty} \mu_1(A_{i,1}(x)).$$

By applying the Monotone convergence theorem on

$$F_k(x) = \sum_{i=1}^{k} \mu_1(A_{i,1}(x))$$

we have that $\mu_1 \times \mu_2$ is countably additive. □

Lecture 7 Product Measure Revisited

7.1 Finishing Touches on Inequality Proof

Recall how previously we proved that, for $p > 1$,

$$|a + b|^p - |b|^p \le \varepsilon |b|^p + C_\varepsilon |a|^p$$

when $|a + b| \ge |b|$ (otherwise it is trivial).

Now consider $|a + b| < |b|$ instead. Then by convexity

$$|b|^p = |b + a - a|^p = \left|\lambda \frac{b + a}{\lambda} - (1 - \lambda)\frac{a}{1 - \lambda}\right|^p \le \lambda \left|\frac{b + a}{\lambda}\right|^p + (1 - \lambda)|\frac{-a}{1 - \lambda}|^p$$
$$= \lambda^{1-p}|b + a|^p + (1 - \lambda)^{1-p}|a|^p.$$

Now choose λ such that $\lambda^{1-p} = 1+\varepsilon$ and $(1-\lambda)^{1-p} = C_\varepsilon$. Then

$$|b|^p \leq (1+\varepsilon)|b+a|^p + C_\varepsilon |a|^p,$$

which rearranged becomes

$$|b|^p - |b+a|^p \leq \varepsilon|b+a|^p - C_\varepsilon|a|^p < \varepsilon|b|^p + C_\varepsilon|a|^p$$

which the last step being true by assumption of $|a+b| < |b|$. Combining these thereby yield

$$||a+b|^p - |b|^p| \leq \varepsilon|b|^p + C_\varepsilon|a|^p.$$

To finish the proof we started last time we require the following theorem:

Theorem 7.1.1 (Uniqueness of Measures)**.** *If Ω is a set, $\mathcal{A}$ an algebra of subsets of Ω, Σ the smallest σ-algebra containing $\mathcal{A}$. Assuming $\varnothing, \Omega \in A$, then if in addition μ is a σ-finite measure in the sense that there exists a collection of subsets $A_i \in \mathcal{A}$ such that $\mu(A_i) < \infty$ and*

$$\bigcup_{i=1}^{\infty} A_i = \Omega,$$

then if μ_2 is a measure such that $\mu_1(A) = \mu_2(A)$ on all $A \in \mathcal{A}$, then $\mu_1 = \mu_2$ on Σ.

Proof continued. Let $\mathcal{A}$ be the set of finite unions of rectangles. That is to say, $A \in \mathcal{A}$ means that $A = A_1 \times A_2$ for $A_1 \in \Omega_1$ and $A_2 \in \Omega_2$, including $\Omega_1 \times \Omega_2$ and $\varnothing$.

Then by construction this is clearly closed under unions, but for it to be an algebra we also require differences to be in there. To see this, first note that

$$(A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \cap A_2) \times (B_1 \cap B_2)$$

meaning that intersections are in there, and then

$$(A_1 \times B_1) - (A_2 \times B_2) = ((A_1 - A_2) \times B_1) \cup ((A_1 \cap A_2) \times (B_1 - B_2)),$$

so differences are in there, and we have an algebra.

Next we verify the strong σ-finiteness: since by assumption there exist $A_i \in \Omega_1$ such that $\mu_1(A_i) < \infty$ and

$$\bigcup_{i=1}^{\infty} A_i = \Omega_1$$

and $B_j \in \Omega$ such that $\mu_2(B_j) < \infty$ and

$$\bigcup_{j=1}^{\infty} B_j = \Omega_2,$$

the collection $A_i \times B_j$ is countable, contains $\Omega_1 \times \Omega_2$, and

$$(\mu_1 \times \mu_2)(A_i \times B_j) = \mu_1(A_i)\mu_2(B_j) < \infty$$

meaning that the measure defined by the integrals is σ-finite in ths stronger sense.

Moreover the measure defined by the two integrals coincides with the above on $\mathcal{A}$, and since $\Sigma_1 \times \Sigma_2$ is the smallest σ-algebra containing $\mathcal{A}$, by the uniqueness of measure theorem the two integrals are indeed equal. □

Corollary 7.1.2. *Product measures commute and are associative.*

Theorem 7.1.3 (Fubini's theorem). *Let $(\Omega_i, \Sigma_i, \mu_i)$ be σ-finite measure spaces for $i = 1, 2$. Let f be $\Sigma_1 \times \Sigma_2$-measurable on $\Omega_1 \times \Omega_2$. If $f \geq 0$, then the following are equal:*

(i) $\displaystyle\int_{\Omega_1 \times \Omega_2} f(x, y)\, (\mu_2 \times \mu_1)(dx\, dx),$

(ii) $\displaystyle\int_{\Omega_1} \int_{\Omega_2} f(x, y)\, \mu_2(dy)\, \mu_1(dx),$

(iii) $\displaystyle\int_{\Omega_2} \int_{\Omega_1} f(x, y)\, \mu_1(dx)\, \mu_2(dy).$

Theorem 7.1.4 (Layer cake representation). *Let ν be a measure on Borel sets of $[0, \infty)$ such that $\varphi(t) = \nu([0, t))$ is finite for all $t > 0$. Let (Ω, Σ, μ) be a measure space and $f \geq 0$ be a Σ-measurable function on Ω. Then*

$$\int_\Omega \varphi(f(x))\, \mu(dx) = \int_0^\infty \mu(\{\, x \mid f(x) > t \,\})\, \nu(dt).$$

Lecture 8 Layer Cake and Egoroff

8.1 Layer Cake Theorem

Proof. We have

$$\int_0^\infty \mu\{\, x \mid f(x) > t \,\}\, \nu(dt) = \int_\Omega \int_0^\infty \chi_{\{\, f>t \,\}}(x)\, \mu(dx)\, \nu(dt)$$

and so applying Fubini's theorem we have

$$\int_\Omega \int_0^\infty \chi_{\{\, f>t \,\}}(x)\, \nu(dt)\, \mu(dx)$$

and since only $0 < t < f(x)$ will contribute to the inner integral,

$$\int_0^\infty \chi_{\{\, f>t \,\}}(x)\, \nu(dt) = \int_0^{f(x)} \nu(dt) = \phi(f(x))$$

and therefore

$$\int_0^\infty \mu\{\, x \mid f(x) > t \,\}\, \nu(dt) = \int_\Omega \phi(f(x))\, \mu(dx).$$

It remains to show that $\mu \times \nu$ is measurable, so that we can in fact apply Fubini's theorem as we did. We check that $\chi_{\{\, f>t \,\}}(x)$ is jointly measurable as a function of (x, t), i.e. $\Sigma \times B^1$-measurability, where B^1 is Borel σ-algebra on R^+.

We compute level sets $F_s = \{\, (x, t) \,|\, \chi_{\{\, f>t \,\}}(x) > s \,\}$. If $s > 1$, then $F_s \neq \varnothing$, and if $0 < s \leq 1$, then $F_s = G = \{\, (x, t) \mid 0 \leq t < f(x) \,\}$.

We need to show that G is measurable, and since

$$G = \bigcup_{r \in \mathbb{Q}} S_f(r) \times [0, r],$$

where $S_f(r) = \{\, x \mid f(x) > r \,\}$. THen G is a countable union of measurable sets and thereby itself measurable, and we are done. ☐

Theorem 8.1.1 (Egoroff's theorem)**.** *Let (Ω, Σ, μ) be a measure space with $\mu(\Omega) < \infty$, and let $f, f_1, f_2, \ldots$ be Σ-measurable on Ω such that $f_j \to f$ almost everywhere on Ω. Then there exists for every $\varepsilon > 0$ some $A_\varepsilon \in \Omega$ such that $\mu(A_\varepsilon) > \mu(\Omega) - \varepsilon$ and $f_j \to f$ uniformly on A_ε.*

In other words we can always remove just a tiny bit of Ω and have uniform convergence.

Proof. Choose $\delta > 0$. Since $f_j \to f$ almost everywhere, there exists some natural number $M(\delta, x)$ such that $|f_j(x) - f(x)| < \delta$ for all $j \geq M(\delta, x)$.

For $N \in \mathbb{N}$, define

$$S(\delta, N) = \{\, x \mid M(\delta, x) \leq N \,\}$$

and

$$S(\delta) = \bigcup_N S(\delta, N).$$

Then

$$S(\delta, N) = \bigcup_{m=1}^{N} \bigcap_{j>m} B_j$$

where $B_j = \{\, x \mid |f_j(x) - f(x)| < \delta \,\}$, meaning that $S(\delta, N)$ is measurable. Moreover almost every x belongs to some $S(\delta, N)$, and wherefore $\mu(S(\delta)) = \mu(\Omega)$.

Next consider a sequence $\delta_1 > \delta_2 > \ldots$, with $\delta_i \to 0$. Let N_j be such that

$$\mu(S(\delta_j, N_j)) > \mu(\Omega) - 2^{-j}\varepsilon$$

and set

$$A_\varepsilon = \bigcap_j S(\delta_j, N_j).$$

We claim that $f_j \to f$ uniformly on A_ε. This is true since for any given $\delta > 0$ falls between δ_j and δ_{j+1} for some j, and for these N_j is the bound for uniform convergence.

Finally we need to measure A_ε, i.e. estimate $\mu(\Omega - A_\varepsilon)$. We have

$$\Big(\bigcap_j S(\delta_j, N_j)\Big)^c = \bigcup_j (S(\delta_j, N_j))^c$$

and because of this

$$\mu\Big(\bigcup_j S(\delta_j, N_j)^c\Big) \leq \sum_j \mu(S(\delta_j, N_j)^c) = \sum_j 2^{-j\varepsilon} = \varepsilon.$$

Therefore $\mu(A_\varepsilon) \geq \mu(\Omega) - \varepsilon$. □

Lecture 9 Simple Functions

Definition 9.0.1 (Simple function)**.** A ***simple function*** on a measure space (Ω, Σ, μ) is a function of the form

$$f(x) = \sum_{j=1}^{N} c_j \chi_{A_j}(x)$$

for $c_j \in \mathbb{R}$ and A_j are μ-measurable sets.

Note that we can choose/decompose A_j in such a way that they are all disjoint, and c_j are distinct.

These are somehow nice because we can describe the integral of them in very simple terms:

$$\int_\Omega f(x)\mu(dx) = \sum_{j=1}^{N} c_j \mu(A_j).$$

Moreover we can, if we so desire, define

$$\int_\Omega f(x)\mu(dx) = \sup_{\substack{0\le g\le f\\ g \text{ simple}}} \int_\Omega g(x)\mu(dx)$$

for μ-measurable functions $f \geq 0$.

Furthermore we can approximate $f \geq 0$ very well by sequences of simple functions, simply by looking at intervals in the range and taking pre-images, then subdividing these by halving them to make the approximation finer.

We can do better: if we let $\mathcal{A}$ be an algebra of sets containing Ω and $\varnothing$, and let Σ be the smallest σ-algebra generated by it, then

Definition 9.0.2 (Really simple function)**.** A ***really simple function*** is a function that can be written as

$$f(x) = \sum_{k=1}^{N} c_k \chi_{A_k}(x),$$

where $c_k \in \mathbb{R}$ and $A_k \in \mathcal{A}$.

Example 9.0.3. Take for instance $\Omega = \mathbb{R}^n$ and let $\mathcal{A}$ be a set of finite unions of half-open rectangles, i.e.

$$\mathcal{A} \ni A = \{\, x \in \mathbb{R}^n \mid a_i < x_i \leq b_i, 1 \leq i \leq n \,\}.$$

▲

Theorem 9.0.4. *Let (Ω, Σ, μ) be a measure space with Σ generated by an algebra $\mathcal{A}$. Assume Ω is σ-finite in the strong sense (relative to $\mathcal{A}$). Then for every $\varepsilon > 0$ and for any summable function f there exists a really simple function h_ε such that*

$$\int_\Omega |f(x) - h_\varepsilon(x)| \, \mu(dx) < \varepsilon.$$

Proof. We can assume that $f \geq 0$, since otherwise we just decompose into positive and negative parts and prove it for them. Next, since for every μ-summable f there exists a simple function f_ε such that

$$\int_\Omega |f(x) - f_\varepsilon(x)| \, \mu(dx) < \varepsilon$$

it is enough to consider simple functions f.

Finally by linearity of integration it suffices to consider f as the characteristic function of a measurable set $C \in \Sigma$, such that $\mu(C) < \infty$.

Next we define $\mathcal{B}$ to be a family of sets $B \in \Sigma$ such that $\mu(B) < \infty$,and for every $\varepsilon > 0$ there exists some $A_\varepsilon \in \mathcal{A}$ such that $\mu(B \triangle A_\varepsilon) < \varepsilon$.

Recall that $X \triangle Y = (X \setminus Y) \cup (Y \setminus X)$ is the symmetric difference.

Note moreover that $\mathcal{A} \subset \mathcal{B}$.

Our goal is to show that $\mathcal{B} = \tilde{\Sigma}$ where $\tilde{\Sigma}$ is the collection of all sets in Σ with finite measure.

First consider the special case $\mu(\Omega) < \infty$. Let B_j, $j = 1, 2, \ldots$ be an increasing family of sets in $\mathcal{B}$ and set $\beta = \bigcup_k B_k$.

Since $\mu(\Omega) < \infty$, we also have $\mu(\beta) < \infty$. Now try to find $A \in \mathcal{A}$ such that $\mu(\beta \triangle A) < \varepsilon$. Set $\sigma_j = \beta - B_j$, and choose j large enough such that $\mu(\sigma_j) < \varepsilon/2$, which is possible since B_j are nested.

By definition of $\mathcal{B}$ we can find $A_j \in \mathcal{A}$ such that $\mu(B_j \triangle A_j) < \varepsilon/2$.

What about $\mu(\beta \triangle A_j)$? First of all $A_j \setminus \beta \subset A_j \setminus B_j$, and thereby $\mu(A_j \setminus \beta) \leq \mu(A_j \setminus B_j)$.

Secondly if we let $X = B_j \setminus A_j$ and $Y = \sigma_j \setminus A_j \subset \sigma_j$, then $\beta \setminus A_j = X \cup Y$.

Therefore

$$\begin{aligned}\mu(\beta \setminus A_j) &\leq \mu(X) + \mu(Y) \leq \mu(X) + \mu(\sigma_j) \\ &= \mu(B_j \setminus A_j) + \mu(\sigma_j) < \mu(B_j \setminus A_j) + \frac{\varepsilon}{2}.\end{aligned}$$

This means that

$$\mu(\beta \triangle A_j) \leq \mu(B_j \setminus A_j) + \mu(A_j \setminus B_j) + \frac{\varepsilon}{2} = \mu(A_j \triangle B_j) + \frac{\varepsilon}{2} = \varepsilon$$

and so $\beta \in \mathcal{B}$. Similarly show that $\bigcap_k B_k \in \mathcal{B}$ if $B_k \in \mathcal{B}$ for a decreasing sequence B_k, making $\mathcal{B}$ a monotone class.

If we also assume temporarily that $\Omega \in \mathcal{A}$, then $\mathcal{B} = \Omega$ by the monotone class theorem. □

Lecture 10 L^p Spaces

10.1 Approximation by Infinitely Differentiable Functions

Theorem 10.1.1. *Let $\Omega \subset \mathbb{R}^n$, μ a Borel measure on Σ, and $\mathcal{A}$ the algebra generated by half-open rectangles. Assume that we have σ-finiteness in the strong sense (relative to $\mathcal{A}$), and that every closed rectangle in Ω has a finite μ-measure.*

If f is a μ-measurable function, then for every $\varepsilon > 0$ there exists a function $g_\varepsilon \in C^\infty(\mathbb{R}^n)$ such that

$$\int_\Omega |f(x) - g_\varepsilon(x)|\, \mu(dx) < \varepsilon.$$

Proof. By considerations from last lecture it is enough to consider f to be the characteristic function of a half-open rectangle H. By assumption $\mu(H) < \infty$.

Next we reduce to a single interval (i.e. one coordinate), and then take the product of the resulting approximations.

In $\mathbb{R}$, therefore, we have $H = (a, b]$. Since Ω is open, there exists a $\delta > 0$ such that $G = [a + \delta, b + \delta]$ such that $\mu(G) < \infty$ by assumption. Let $h_\varepsilon(x) = \phi(x/\varepsilon)$, where

$$\phi(x) = \begin{cases} \exp(-(e^{x/(1-x)} - 1)^{-1}), & \text{if } 0 < x < 1 \\ 0, & \text{if } x \le 0 \\ 1, & \text{if } x \ge 1. \end{cases}$$

Now define

$$g_\varepsilon(x) = \begin{cases} h_\varepsilon(x-a), & \text{if } x \le a+\varepsilon \\ 1 & \text{if } a+\varepsilon < x < b \\ h_\varepsilon(b+\varepsilon-x) & \text{if } x \ge b. \end{cases}$$

This function g_ε has three useful properties:

(*i*) $g_\varepsilon(x) \in C^\infty(\mathbb{R})$,

(*ii*) $g_\varepsilon(x) \to \chi_H(x)$ for every $x \in \mathbb{R}$ as $\varepsilon \to 0$,

(*iii*) $0 \le g_\varepsilon(x) \le \chi_H(x) + \chi_G(x)$ when $\varepsilon < \delta$. □

10.2 L^p Spaces

Definition 10.2.1 (L^p space)**.** Let (Ω, Σ, μ) be a measure space, and fix a real number $1 \le p < \infty$. We then define

$$L^p(\Omega, d\mu) = \{\, f\colon \Omega \to \mathbb{R} \mid f \text{ is } \mu\text{-measurable and } |f|^p \text{ is } \mu\text{-summable}\,\}.$$

We claim that L^p is a linear vector space, meaning that it is closed under multiplication by scalars (which is clear) and closed under addition. To see this, use the fact that $|\cdot|^p$ is convex for $1 \le p < \infty$, whereby $|f(x)+g(x)|^p \le 2^{p-1}(|f(x)|^p + |g(x)|^p)$, and since the right-hand side is summable, so it the left-hand side.

Definition 10.2.2. The norm on $L^p(\Omega, d\mu)$ is defined by

$$\|f\|_p = \Big(\int_\Omega |f(x)|^p \,\mu(dx)\Big)^{1/p},$$

meaning that

(*i*) $\|\lambda f\|_p = |\lambda| \|f\|_p$ for all $\lambda \in \mathbb{R}$,

(*ii*) $\|f\|_p = 0$ implies that $f(x) = 0$ μ-almost everywhere, and

(*iii*) $\|f+g\|_p \le \|f\|_p + \|g\|_p$.

Lecture 11 Jensen's Inequality

We'll note first of all that we neglected $p = \infty$ when we defined L^p spaces last time:

$$L^\infty(\Omega, d\mu) = \{\, f\colon \Omega \to \mathbb{R} \mid f \text{ is } \mu\text{-measurable, } |f(x)| \le k \text{ for some } k \ \mu\text{-a.e.}\,\}$$

with the norm

$$\|f\|_\infty = \inf\{\, k \mid |f(x)| \le k \ \mu\text{-a.e.}\,\}.$$

11.1 Convexity

We'll make a note about convexity: if f is convex on an open set K, then f is continuous on K. Moreover we can construct a support plane—generalising the tangent line—of $f\colon K \to \mathbb{R}$, $K \subset \mathbb{R}^n$, at a point $x \in K$. This is an $\mathbb{R}^{n+1}$ hyperplane that touches the graph at $(x, f(x))$ and never lies above the graph.

In other words,

$$f(y) \geq f(x) + V \cdot (y - x)$$

for every $y \in K$ with $V \in \mathbb{R}^{n+1}$ depending on x.

11.2 Jensen's Inequality

Theorem 11.2.1 (Jensen's inequality)**.** *Let $J\colon \mathbb{R} \to \mathbb{R}$ be a convex function, let $f\colon \Omega \to \mathbb{R}$ be measurable with respect to (Ω, Σ, μ). Assume $\mu(\Omega) < \infty$, and finally suppose $f \in L^1(\Omega, d\mu)$. Then we define the integral average*

$$\langle f \rangle = \frac{1}{\mu(\Omega)} \int_\Omega f(x)\, \mu(dx).$$

With all of these assumptions, then $[J \circ f]_- \in L^1(\Omega, d\mu)$ (meaning that the integral of $J \circ f$ is well-defined) and

$$\langle J \circ f \rangle \geq J(\langle f \rangle).$$

In other words

$$\frac{1}{\mu(\Omega)} \int_\Omega J(f(x))\, \mu(dx) \geq J\Big(\frac{1}{\mu(\Omega)} \int_\Omega f(x) \mu(dx)\Big).$$

If J is strictly convex, then equality holds only if f is a constant function.

Proof. By convexity, there exists a $V \in \mathbb{R}$ such that

$$J(t) \geq J(\langle f \rangle) + V(t - \langle f \rangle)$$

for every $t \in \mathbb{R}$. Thus taking $t = f(x)$ we have

$$(J \circ f)(x) \geq J(\langle f \rangle) + V f(x) - V \langle f \rangle.$$

Now taking absolute values in the left-hand side will preserve the inequality, and then multiplying by -1 we'll flip the inequality, so

$$-|J(f(x))| \leq -J(\langle f \rangle) - V f(x) + V \langle f \rangle$$

and now taking absolute value of the right-hand side will preserve this new inequality, and therefore (using the triangle inequality on the way)

$$-|J(f(x))| \leq |J(\langle f \rangle)| + |V||f(x)| + |V||\langle f \rangle|.$$

Now the e negative absolute value of a function is, in places where the function itself negative, exactly the negative part of the function. Moreover elsewhere the negative part is 0, and in our right-hand side we have the sum of nonnegative terms, and therefore

$$[J \circ f]_-(x) \leq |J(\langle f \rangle)| + |V||f(x)| + |V||\langle f \rangle|$$

which is in L^1 and so our integral is well-defined.

To get Jensen's inequality now, we simply integrate

$$J(f(x)) \geq J(\langle f \rangle) + V(f(x) - \langle f \rangle)$$

yielding

$$\int_\Omega J(f(x))\,\mu(dx) \geq \int_\Omega J(\langle f \rangle)\,\mu(dx) + V\Big(\int_\Omega f(x)\,\mu(dx) - \langle f \rangle \int_\Omega \mu(dx)\Big).$$

Since by definition

$$\int_\Omega f(x)\,\mu(dx) = \langle f \rangle \mu(\Omega)$$

the last term in the right-hand side is zero. We then divide both sides by $\mu(\Omega)$, and we have Jensen's inequality. □

Another handy inequality is this:

Theorem 11.2.2 (Young's inequality)**.** *Suppose $\varphi\colon \mathbb{R} \to \mathbb{R}$ is continuous and strictly increasing for $x \geq 0$, and $\varphi(0) = 0$. If $x = \psi(y) = \varphi^{-1}(y)$ and $a, b > 0$, then*

$$ab \leq \int_0^a \varphi(x)\,dx + \int_0^b \psi(y)\,dy.$$

Example 11.2.3. If $\varphi(x) = x^\alpha$, $\alpha > 0$, then $\psi(y) = y^{1/\alpha}$. Then

$$ab \leq a^{1+\alpha}\frac{1}{1+\alpha} + \frac{1}{1+1/\alpha}b^{1+1/\alpha}$$

and so, taking $p = 1 + \alpha$ and $q = 1 + 1/\alpha$, and noting that $1/p + 1/q = 1$, then

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

▲

Using this inequality we get Holder's inequality for free. If $\|f\|_p = \|g\|_q = 1$, then

$$|fg| \leq \frac{|f|^p}{p} + \frac{|g|^q}{q}$$

and if we integrate both sides we have

$$\int_\Omega |fg|\,d\mu \leq \frac{1}{p}\|f\|_p^p + \frac{1}{q}\|g\|_q^q = \frac{1}{p} + \frac{1}{q} = 1 = \|f\|_p \|g\|_q.$$

If the norms aren't one, we just normalise by dividing by the respective norms, so long as they aren't 0.

Lecture 12 Hölder's Inequality

12.1 Hölder's Inequality

A useful property that we have not yet proved is this:

Proposition 12.1.1. *If $\mu(\Omega) < \infty$, then $\|f\|_\infty = \lim_{p\to\infty} \|f\|_p$.*

Proof. Let $M = \|f\|_\infty$. For $k < M$, $\mu(\{ x \in \Omega \mid |f(x)| > k \}) > 0$. Setting $S = \{ x \in \Omega \mid |f(x)| > k \}$, we have

$$\|f\|_p \geq \Big(\int_S |f|^p \, d\mu \Big)^{1/p} \geq k\mu(S)^{1/p}.$$

As $p \to \infty$, $\mu(S)^{1/p} \to 1$, and so $\|f\|_p \geq k$ for every $k < M$, so

$$\liminf_{p\to\infty} \|f\|_p \geq M.$$

Secondly,

$$\|f\|_p \leq \Big(\int_\Omega M^p \, d\mu \Big)^{1/p} = M\mu(\Omega)^{1/p}$$

which implies that

$$\limsup_{p\to\infty} \|f\|_p \leq M$$

and therefore the limit is $M = \|f\|_p$. □

Theorem 12.1.2 (Hölder's inequality). *Let $f \in L^p(\Omega, d\mu)$ and $g \in L^q(\Omega, d\mu)$ such that $1/p + 1/q = 1$, with $1 \leq p \leq \infty$. Then $fg \in L^q(\Omega, d\mu)$ and*

$$\left| \int_\Omega fg \, d\mu \right| \leq \int_\Omega |f||g| \, d\mu \leq \|f\|_p \|g\|_q.$$

Equality in the first inequality holds if and only of $fg = |f||g|$ μ-almost everywhere.

If $f \neq 0$, equality in the second inequality holds if and only if thjere exists some $\gamma \in \mathbb{R}$ such that

(*i*) *if $1 < p < \infty$, then $|g(x)| = \lambda |f(x)|^{p-1}$ μ-almost everywhere;*

(*ii*) *if $p = 1$, then $|g(x)| \leq \lambda$ for μ-almost everywhere x and $|g(x)| = \lambda$ when $f(x) \neq 0$; and*

(*iii*) *if $p = \infty$, then $|f(x)| \leq \lambda$ for μ-almost everywhere x and $|f(x)| = \lambda$ when $g(x) \neq 0$.*

Proof. Assume without loss of generality that $f, g \geq 0$, since we're taking absolute values anyway. Let $1 < p, q < \infty$. Set $A = \{ x \mid f(x) > 0 \} \subset \Omega$, so that $B = \Omega \setminus A = \{ x \mid g(x) = 0 \}$. This is because if we work over A, then we can divide by g.

Note that

$$\int_\Omega f^p \, d\mu = \int_A f^p \, d\mu + \int_B f^p \, d\mu$$

and

$$\int_\Omega g^q \, d\mu = \int_A g^q \, d\mu,$$

meaning that

$$\int_\Omega fg \, d\mu = \int_A fg \, d\mu.$$

For this reason we may assume $\Omega = A$. Now define a measure ν on A by $\nu(dx) = g(x)^q\,\mu(dx)$. Set $F(x) = f(x)q^{-1/p}$ on A.

Consider
$$\langle F\rangle_\nu = \frac{1}{\nu(\Omega)}\int_\Omega F(x)\,\nu(dx).$$
We have
$$\nu(\Omega) = \int_\Omega \nu(dx) = \int_\Omega g^q\,d\mu = \|g\|_q^q$$
which is finite by assumption. We also have
$$\begin{aligned}\int_\Omega F(x)\,\nu(dx) &= \int_\Omega f(x)g(x)^q g(x)^{-q/p} g(x)^q\,d\mu \\ &= \int_\Omega f(x)g(x)^{q(1-1/p)}\,d\mu = \int_\Omega f(x)g(x)^{q/q}\,d\mu = \int_\Omega fg\,d\mu\end{aligned}$$
since $1/p + 1/q = 1$. Thus
$$\langle F\rangle_\nu = \frac{\int_\Omega fq\,d\mu}{\int_\Omega g^q\,d\mu}.$$
On the other hand taking $J(t) = |t|^p$,
$$\int_\Omega J(F)\,d\nu = \int_\Omega |F|^p\,d\nu = \int_\Omega f(x)^p g(x)^{-q} g(x)^q\,\mu(dx) = \int_\Omega f(x)^p\,d\mu.$$
If we now use Jensen with J on F with respect to ν, then
$$\langle J(F)\rangle \geq J(\langle F\rangle),$$
or in other words
$$\frac{\int_\Omega f^p\,d\mu}{\int_\Omega g^q\,d\mu} \geq \frac{\left(\int_\Omega fg\,d\mu\right)^p}{\left(\int_\Omega g^q\,d\mu\right)^p}.$$
Taking $1/p$th powers of both sides, we have
$$\frac{\left(\int_\Omega f^p\,d\mu\right)^{1/p}}{\left(\int_\Omega g^q\,d\mu\right)^{1/p}} \geq \frac{\int_\Omega fg\,d\mu}{\int_\Omega g^q\,d\mu}.$$
Rearranging, we get the inequality:
$$\left(\int_\Omega f^p\,d\mu\right)^{1/p}\left(\int_\Omega g^q\,d\mu\right)^{1-1/p} \geq \int_\Omega fg\,d\mu$$
which is what we want since $1 - 1/p = 1/q$. □

12.2 Completeness of L^p Spaces

It so happens that spaces thus defined are complete—meaning that Cauchy sequences converge to elements in the space.

Theorem 12.2.1. *Let $1 \leq p \leq \infty$ and let $\{ f_j \}$ be a Cauchy sequence in $L^p(\Omega, d\mu)$. Then*

(i) *There exists a unique function $f \in L^p(\Omega, d\mu)$ such that $\|f_i - f\|_p \to 0$ as $i \to \infty$.*

(ii) *There exists a subsequence $\{ f_{i_k} \}$ and a function $F(x) \geq 0$ in L^p such that $|f_{i_k}(x)| \leq F(x)$ for every k μ-almost everywhere x, and*

$$\lim_{k\to\infty} f_{i_k}(x) = f(x)$$

μ-almost everywhere.

Proof. It is enough to show (i) for some subsequence $\{ f_{i_k} \}$ because

$$\|f_i - f\|_p \leq \|f_i - f_{i_k}\|_p + \|f_{i_k} - f\|_p$$

by the triangle inequality. The first one goes to 0 by the Cauchy property, and the second goes to 0 if $f_{i_k} \to f$. We construct $\{ f_{i_k} \}$ as follows: Pick i_1 such that $\|f_{i_k} - f_n\|_p \leq 1/2$ for all $n \geq i_1$, which is possible since the sequence if Cauchy. Next pick i_2 such that $\|f_{i_2} - f_n\|_p \leq 1/4$ for all $n \geq i_2$, and so on, so that $\|f_{i_k} - f_{i_{k+1}}\|_p < 2^{-k}$ and $i_1 \leq i_2 \leq \ldots$.

Define

$$F_\ell(x) = |f_{i_1}(x)| + \sum_{k=1}^{\ell} |f_{i_k}(x) - f_{i_{k+1}}(x)|$$

motivated by

$$f_{i_{k+1}}(x) = f_{i_1}(x) + \sum_{\ell=1}^{k} f_{i_{\ell+1}}(x) - f_{i_\ell}(x).$$

Now by the triangle inequality

$$\|F_\ell\|_p \leq \|f_{i_1}\|_p + \sum_{k=1}^{\infty} \|f_{i_k} - f_{i_{k+1}}\|_p \leq \|f_{i_1}\|_p + \sum_{k=1}^{\infty} 2^{-k} \leq \|f_{i_1}\|_p + 1.$$

Thus there exists some $F(x)$ such that $F_\ell(x) \to F(x)$ μ-almost everywhere since $F_\ell(x)$ is increasing and bounded.

Now by the monotone convergence theorem

$$\lim_{\ell\to\infty} \int_\Omega |F_\ell|^p \, d\mu = \int_\Omega |F|^p \, d\mu$$

and so $F \in L^p(\Omega, d\mu)$ and in particular it is finite μ-almost everywhere. □

Lecture 13 Minkowski's Inequality

We start by finishing up the proof of the completeness of L^p.

Proof continued. We considered partial sums of the series

$$\sum_{k=1}^{\infty} f_{i_{k+1}}(x) - f_{i_k}(x)$$

and we proved that this series converges absolutely almost everywhere. Therefore it converges. In other words the sequence $\{ f_{i_k} \}$ converges absolutely, so it converges, and so there exists some f such that $f_{i_k}(x) \to f(x)$ μ-almost everywhere.

We now know that $|f_{i_k} - f|^p(x) \to 0$ μ-almost everywhere. To see that the integral does the same, we bound the integrand by a summable function, which we already have in hand:

$$|f_{i_k}(x) - f(x)| \le F(x) + |f(x)| \in L^p(\Omega, d\mu),$$

and so by the Dominated convergence theorem $\|f_{i_k} - f\|_p \to 0$. □

13.1 Minkowski's Inequality

Theorem 13.1.1 (Minkowski's inequality)**.** *Suppose* (Ω, Σ_1, μ) *and* (Γ, Σ_2, ν) *are σ-finite measure spaces. Suppose further that f is a $\mu \times \nu$-measurable function on $\Omega \times \Gamma$ that is nonnegative. Let $1 \le p < \infty$. Then*

$$\int_\Gamma \Big(\int_\Omega f(x,y)^p \, \mu(dx) \Big)^{1/p} \nu(dy) \ge \Big(\int_\Omega \Big(\int_\Gamma f(x,y) \, \nu(dy) \Big)^p \mu(dx) \Big)^{1/p}.$$

A special case is

$$\|f+g\|_p \le \|f\|_p + \|g\|_p,$$

the triangle inequality for L^p norm.

Proof. First we establish how the triangle inequality follows from the multiplicative inequality. The trick is to choose one of the two measures quite cleverly, namely let $\Gamma = \{1, 2\}$ and ν the counting measure on this set. Then let $F(x,1) = |f(x)|$ and let $F(x,2) = |g(x)|$.

Thus

$$\int_\Gamma F(x,y) \, \nu(dy) = F(x,1) + F(x,2) = |f(x)| + |g(x)|,$$

which is $f(x) + g(x)$ if we assume f and g nonnegative. Applying Minkowski's inequality to this we get

$$\begin{aligned} \Big((f(x)+g(x))^p \, \mu(dx) \Big)^{1/p} &\le \int_\Gamma \underbrace{\Big(\int_\Omega F(x,y)^p \, \mu(dx) \Big)^{1/p}}_{=G(y)} \nu(dy) = G(1) + G(2) \\ &= \Big(\int_\Omega f(x)^p \, \mu(dx) \Big)^{1/p} + \Big(\int_\Omega g(x)^p \, \mu(dx) \Big)^{1/p}, \end{aligned}$$

so in other words $\|f+g\|_p \le \|f\|_p + \|g\|_p$.

Now to prove the actual inequality at hand, note that by Fubini's theorem the functions

$$\int_\Omega f(x,y)^p \, \mu(dx) \quad \text{and} \quad \int_\Gamma f(x,y) \, \nu(dy)$$

are both measurable. We can assume that $f(x,y) > 0$ on a set of positive $\mu \times \nu$-measure, and we can also assume that f is finite, otherwise we truncate and use monotone convergence.

Therefore we may assume in what follows that

$$\int_\Gamma \Big(\int_\Omega f(x,y)^p \, \mu(dx) \Big)^{1/p} \nu(dy)$$

is finite. □

Lecture 14 Weak Convergence

We start by finishing up the proof of Minkowski's inequality.

Proof. Let

$$H(x) = \int_\Gamma F(x, y)\, \nu(dy).$$

Then the right-hand side of Minkowski is

$$\begin{aligned}\Big(\int_\Omega H(x)^p\, \mu(dx)\Big)^{1/p} &= \Big(\int_\Omega H(x)H(x)^{p-1}\, \mu(dx)\Big)^{1/p} \\ &= \Big(\int_\Omega \Big(\int_\Gamma f(x, y)\, \mu(dy)\Big) H(x)^{p-1}\, \mu(dx)\Big)^{1/p} \\ &= \Big(\int_\Omega \int_\Gamma f(x, y)H(x)^{p-1}\, \nu(dy)\, \mu(dx)\Big)^{1/p} \\ &= \Big(\int_\Gamma \Big(\int_\Omega f(x, y)H(x)^{p-1}\, \mu(dx)\Big)\, \nu(dy)\Big)^{1/p}\end{aligned}$$

which we can bound by

$$\Big(\int_\Gamma \Big(\Big(\int_\Omega f(x, y)^p\, \mu(dx)\Big)^{1/p} \Big(\int_\Omega H(x)^{(p-1)p/(p-1)}\, \mu(dx)\Big)^{(p-1)/p} \nu(dy)\Big)^{1/p}$$

using Hölder. Raising both sides of the inequality to p and dividing through by the inner integral, we have

$$\Big(\int_\Omega H(x)^p\, \mu(dx)\Big)^{1-(p-1)/q} \le \int_\Gamma \Big(\int_\Omega f(x, y)^p\, \mu(dx)\Big)^{1/p} \nu(dy)$$

in which the power on the left-hand side is $1 - 1/q = 1/p$, which gives of Minkowski. □

14.1 Dual Space

Definition 14.1.1 (Linear functional)**.** A ***linear functional*** on $L^p(\Omega, d\mu)$ is a map $L\colon L(\Omega, d\mu) \to \mathbb{R}$ such that

$$L(a_1 f_1 + a_2 f_2) = a_1 L(f_1) + a_2 L(f_2)$$

for every $a_1, a_2 \in \mathbb{R}$ and every $f_1, f_2 \in L^p(\Omega, d\mu)$.

A linear functional L is ***continuous*** if $f_i \to f$ strongly (in norm) in $L^p(\Omega, d\mu)$ implies $L(f_i) \to L(f)$.

A linear functional L is ***bounded*** if there exists some $k \ge 0$ such that $|L(f)| \le k\|f\|_p$.

Remark 14.1.2. Note that for linear operators, continuity is equivalent to boundedness (this is true in any normed space).

Definition 14.1.3 (Dual space)**.** The set of continuous linear functionals on $L^p(\Omega, d\mu)$ is the ***dual spacedual space*** $L^p(\Omega, d\mu)^*$. The space can be equipped with the norm

$$\|L\| = \sup_{\|f\|_p \le 1} |L(f)|.$$

Remark 14.1.4. This is indeed a norm. All but the triangle inequality are obvious, and the triangle inequality isn't very hard either:

$$\begin{aligned}\|L_1 + L_2\| &= \sup_{\|f\|_p \le 1} |(L_1 + L_2)(f)| = \sup_{\|f\|_p \le 1} |L_1(f) + L_2(f)| \\ &\le \sup_{\|f\|_p \le 1} |L_1(f)| + \sup_{\|f\|_p \le 1} |L_2(f)| = \|L_1\| + \|L_2\|.\end{aligned}$$

Definition 14.1.5 (Weak convergence)**.** We say that a sequence $\{ f_i \} \subset L^p(\Omega, d\mu)$ ***converges weakly*** to $f \in L^p(\Omega, d\mu)$ if $L(f_i) \to L(f)$ for all $L \in L^p(\Omega, d\mu)^*$. We write $f_i \rightharpoonup f$.

It is worth noting that weak convergence, as the name suggests, does imply strong convergence, specifically because $|L(f_i - f)| \le \|L\| \|f_i - f\|_p$.

We will give three generic type of situations where weak convergence fails to imply strong convergence. The first case is oscillations: suppose $f_i(x) = \sin(ix)$ on $0 \le x \le 1$.

The second situation is when $f_i(x)$ becomes very large at a certain location. Say $f_i(x) = i^{1/p} g(ix)$, with g a fixed function in $L^p(\Omega, d\mu)$.

Finally the sequence f_i might leak to infinity, say $f_i(x) = g(x + i)$ for some $g \in L^p(\Omega, d\mu)$.

In all of these cases, $f_i \rightharpoonup 0$, but none of them converge strongly.

14.2 Characterising $L^p(\Omega, d\mu)^*$

From Hölder we have $L^q(\Omega, d\mu) \subset L^p(\Omega, d\mu)^*$ for $1/p + 1/q = 1$ since we can view

$$L_g(f) = \int_\Omega gf \, d\mu$$

for all $g \in L^q(\Omega)$. The natural follow up question is whether $L^q(\Omega, d\mu) = L^p(\Omega, d\mu)^*$. The answer is yes if $1 \le p < \infty$, and no if $p = \infty$.

We will make use of the following theorem, though we will not prove it:

Theorem 14.2.1 (Hanner's inequality)**.** *Let $f, g \in L^p(\Omega, d\mu)$. If $1 \le p \le 2$, then*

$$\|f + g\|_p^p - \|f - g\|_p^p \ge (\|f\|_p + \|g\|_p)^p + |\|f\|_p - \|g\|_p|^p$$

and

$$(\|f + g\|_p + \|f - g\|_p)^p + |\|f + g\|_p - \|f - g\|_p|^p \le 2^p(\|f\|_p^p + \|g\|_p^p).$$

If $2 \le p < \infty$, the inequalities reverse.

Lemma 14.2.2 (Projection on convex set)**.** *Let $1 < p < \infty$, and K a convex set in $L^p(\Omega, d\mu)$ (meaning that for all $f_1, f_2 \in K$ and $\lambda \in [0, 1]$, $\lambda f_1 + (1 - \lambda) f_2 \in K$). For any $f \in L^p(\Omega, d\mu)$ and $f \notin K$, define the distance*

$$D(f, k) = \int_{g \in K} \|f - g\|_p.$$

Then there exists a function $h \in K$ such that $D(f, k) = \|f - k\|_p$ (i.e. the infimum is attained). Moreover

$$\int_\Omega (f - h)(f - h)|f - h|^{p-2} \, d\mu \le 0$$

for all $g \in K$ and all $f \in L^p(\Omega, d\mu)$, and $f \notin K$.

Lecture 15 Projection on a Convex Set

15.1 Proof of the Lemma

We start by proving the lemma stated at the end of last lecture.

Proof. Let $1 < p \leq 2$, and assume without loss of generality that $f = 0$ (if not we could just shift).

Let $\{ h_j \} \subset K$ be a minimising sequence, i.e. $\|h_j\|_p \to D = \|h\|_p$. We want to show first that there exists some h such that $h_j \to h$, and secondly that $h \in K$. For the first one, we need to show that $\{ h_j \}$ is Cauchy.

Note that $\|h_j + h_k\|_p \leq \|h_j\|_p + \|h_k\|_p \to 2D$ as $j, k \to \infty$. Moreover $1/2(h_j + h_k) \in K$ since K is a convex set, and D is the infimum of all norms on K, so $\|h_j + h_k\|_p \geq 2D$. Therefore

$$\limsup \|h_j + h_k\|_p \leq 2D \leq \liminf \|h_j + h_k\|_p,$$

so $\|h_j + h_k\|_p \to 2D$.

We now set up Hanner's inequality on these:

$$(\|h_j + h_k\|_p + \|h_j - h_k\|_p)^p + |\|h_j + h_k\|_p - \|h_j - h_k\|_p|^p \leq 2^p(\|h_j\|_p^p + \|h_k\|_p^p).$$

We know that the right-hand side goes to $2^{p+1}D^p$, and that the $h_j + h_k$ parts alone in the left-hand side make the left-hand side go to the same, so in some sense there's 'not enough room' for $\|h_j - h_k\|_p \not\to 0$. For the same of argument, assume that it doesn't go to 0.

In other words, for infinitely many j and k, there exists $b > 0$ such that $\|h_j - h_k\|_p \geq b$. Then we would have

$$|2D + b|^p + |2D - b|^p \leq 2^{p+1}D^p,$$

wherein the two absolute values are increasing in b, so as we let b go to zero, the left-hand side decreases.

Let $u(x) = |x|^p$, and take $z = 2D + x$ and $y = 2D - x$. Then since u is strictly convex for $x \neq 0$,

$$u\Big(\frac{1}{2}z + \frac{1}{2}y\Big) < \frac{1}{2}u(z) + \frac{1}{2}u(y)$$

whereby

$$\frac{1}{2^p}|z + y|^p < \frac{1}{2}(|z|^p + |y|^p).$$

Rearranging this we get

$$|z|^p + |y|^p > \frac{1}{2^{p-1}}|z + y|^p,$$

so

$$|2D + x|^p + |2D - x|^p > \frac{1}{2^{p-1}}(4D^p) = \frac{1}{2^{p-1}}2^{2p}D^p = 2^{2p-p+1}D^p = 2^{p+1}D^p$$

Now if $b > 0$ we have

$$2^{p+1}D^p < |2D + b|^p + |2D - b|^p \leq 2^{p+1}D^p,$$

i.e. $0 < 0$, which is impossible. Hence $b = 0$, so $\|h_j - h_k\|_p \to 0$ as $j, k \to \infty$, so $\{ h_j \}$ is Cauchy, and therefore converges to $h \in L^p(\Omega)$ since the space is complete. Moreover $h \in K$ since K is norm closed, implying that $D = \|f - h\|_p$.

For the second part of the lemma, namely the inequality, given any $g \in K$, let $g_t = (1-t)h + tg$ for $t \in [0, 1]$. Since K is a convex set, $g_t \in K$. Now consider $F(t) = \|f - g^t\|_p^p$. From the previous part, $F(0) = D^p$. For other t, $F(t) \geq D^p$ for $0 < t \leq 1$, again by the above. Consider the derivative

$$\begin{aligned} F'(t) &= \frac{d}{dt} \int_\Omega |f - (1-t)h - tg|^p \, d\mu \\ &= \int_\Omega p|f - (1-t)h - tg|^{p-1} \frac{d}{dt} |f - (1-t)h - tg| \, d\mu \\ &= \int_\Omega p|f - (1-t)h - tg|^{p-1} \frac{f - (1-t)h - tg}{|f - (1-t)h - tg|} (h - g) \, d\mu \\ &= p \int_\Omega |f - (1-t)h - tg|^{p-2} (f - (1-t)h - tg)(h - g) \, d\mu, \end{aligned}$$

which implies that

$$F'(0) = p \int_\Omega |f - h|^{p-2} (f - h)(h - g) \, d\mu,$$

and since $F(t) - F(0) \geq 0$ for all $0 < t \leq 1$, $F'(0) \geq 0$, so

$$\int_\Omega |f - h|^{p-2} (f - h)(g - h) \, d\mu \leq 0.$$

For $2 < p < \infty$, we do almost the same thing, but using the opposite version of Hanner's inequality. □

Theorem 15.1.1. *Suppose $f \in L^p(\Omega)$ satisfies $L(f) = 0$ for all $L \in L^p(\Omega)^*$. Then $f = 0$.*

In other words, $f_j \rightharpoonup f$ and $f_j \rightharpoonup g$ implies $f = g$—i.e. weak limits are unique.

Proof. If $1 < p < \infty$, then

$$g(x) = \begin{cases} |f(x)|^{p-2} f(x) & \text{if } f(x) \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

(The 0 part is of course to avoid potential division by 0.) Now $f \in L^p(\Omega)$ implies $g \in L^q(\Omega)$, since

$$||f(x)|^{p-2} f(x)|^q = |f(x)|^{(p-1)q} = |f(x)|^p.$$

Moreover,

$$\int_\Omega gf \, d\mu = \|f\|_p^p$$

since $gf = |f|^{p-2} f \cdot f = |f|^p$. Next, the functional

$$L(g) = \int_\Omega gh \, d\mu$$

is a continuous linear functional for $h \in L^p(\Omega)$. Note that it is continuous because it is bounded, and it is bounded because of Hölder's inequality. If we now take $h = f$, then

$$L(f) = \int_\Omega gf \, d\mu = \|f\|_p^p = 0.$$

Therefore $f = 0$ almost everywhere. The cases of $p = 1$ and $p = \infty$ are straightforward. □

Lecture 16 More on Weak Convergence

16.1 Upgrading Weak Convergence

The goal in the immediate future is to explore the difference between $\|f_j\|_p$ and $\|f\|_p$ if we know that $f_j \rightharpoonup f$.

Theorem 16.1.1. *If $1 \le p \le \infty$ and if $f_j \rightharpoonup f$, then*

(i) $\liminf_{j\to\infty} \|f_j\|_p \ge \|f\|_p$ *(if $p = \infty$, we also need to assume that μ is σ-finite).*

(ii) If $1 < p < \infty$ and if $\|f_j\|_p \to \|f\|_p$, then $f_j \to f$ in $L^p(\Omega)$.

Proof. Let $1 \le p < \infty$, and consider

$$L(h) = \int_\Omega gh \, d\mu$$

where $g(x) = |f(x)|^{p-2} f(x)$. Note that $L(f) = \|f\|_p^p$. Then

$$L(f_j) = \int_\Omega g f_j \, d\mu \le \|g\|_q \|f_j\|_p,$$

meaning that

$$\lim_{j\to\infty} L(f_j) \le \|g\|_q \liminf_{j\to\infty} \|f_j\|_p.$$

But $L(f_j) \to L(f)$ by weak convergence, so $\|f\|_p^p \le \|g\|_q \liminf \|f_j\|_p$, and $\|g\|_q = \|f\|_p^{p-1}$, which in turn implies that $\|f\|_p^p \le \|f\|_p^{p-1} \liminf \|f_j\|_p$. Therefore the first claim is true.

Next, if $p = \infty$, assume $\|f\|_\infty = a > 0$ (the alternative, that $a = 0$, is trivial). Let $A_\varepsilon = \{ x \in \Omega \mid |f(x)| > a - \varepsilon \}$. By σ-finiteness, there exist B_k with $\mu(B_k) < \infty$ and $A_\varepsilon \cap B_k$ increasing to A_ε. Define

$$g_{k,\varepsilon}(x) = \begin{cases} f(x)/|f(x)| & \text{if } x \in A_\varepsilon \cap B_k \\ 0 & \text{otherwise.} \end{cases}$$

Now

$$\int_\Omega g_{k,\varepsilon} \, d\mu = \int_{A_\varepsilon \cap B_k} \frac{f(x) f_j(x)}{|f(x)|} \, d\mu \le \|f_j\|_\infty \int_{A_\varepsilon \cap B_k} \frac{|f(x)|}{|f(x)|} \, d\mu = \|f_j\|_\infty \mu(A_\varepsilon \cap B_k),$$

by Hölder's inequality. Since

$$L(h) = \int_\Omega g_{k,\varepsilon} h \, d\mu$$

is a linear operator, by weak convergence we have

$$\lim_{j\to\infty}\int_\Omega g_{k,\varepsilon} f_j\, d\mu = \int_\Omega g_{k,\varepsilon} f\, d\mu = \int_{A_\varepsilon\cap B_k} |f(x)|\, d\mu \le \liminf_{j\to\infty} \|f_j\|_\infty \mu(A_\varepsilon \cap B_k)$$

as $j \to \infty$.

But

$$\int_{A_\varepsilon\cap B_k} |f(x)|\, d\mu \ge (a-\varepsilon)\mu(A_\varepsilon \cap B_k),$$

so $\liminf\|f_j\|_\infty \ge a-\varepsilon$ for all $\varepsilon > 0$, and so $\liminf\|f_j\|_\infty \ge a = \|f\|_\infty$.

To prove the second assertion, for $1 < p < \infty$ we note that $\|f_j\|_p \to \|f\|_p$ implies that $\|f_j + f\|_p \to 2\|f\|_p$.

Indeed $f_j + f \rightharpoonup 2f$ and so by the first part of the theorem, $\liminf\|f_j+f\|_p \ge 2\|f\|_p$ and $\|f_j + f\|_p \le \|f_j\| + \|f\|_p$, so $\limsup\|f_j + g\|_p \le 2\|f\|_p$, and so in all

$$\lim_{j\to\infty} \|f_j + f\|_p = 2\|f\|_p.$$

If we restrict ourselves to $1 < p \le 2$, then by Hanner's inequality with $g = f_j$ and $f = f$, letting $A_j = \|f + f_j\|_p$ and $B_j = \|f - f_j\|_p$, we have

$$\limsup_{j\to\infty}((A_j + B_j)^p + |A_j - B_j|^p) \le 2^{p+1}\|f\|_p^p.$$

We would like to prove that $B_j \to 0$. One way to do this is to follow along with what we did last time in the proof of projection on convex sets. Another argument is to note that from convexity of $|A + x|^p$ we get a lower bound

$$2A_j^p \le (A_j + B_j)^p + |B_j - A_j|^p \le 2^p\|f\|_p^p + 2^p\|f_j\|_p^p.$$

Note that $A_j^p \to 2^p\|f\|_p^p$, and so if we divide by A_j and subtract 2 we have

$$0 \le (1 + B_j/A_j)^p + |B_j/A_j - 1|^p - 2 \le \frac{2^p\|f\|_p^p + 2^p\|f_j\|_p^p}{A_j^p} - 2. \qquad (16.1.1)$$

Let $t = B_j/A_j$ and consider for $t < 1$ the function $f(t) = (1 + t)^p + |1 - t|^p - 2$. We have $f'(t) = p((1+t)^{p-1} - (1-t)^{p-1})$, which is concave, and so $f'(t) \ge 2^{p-1}t$. Integrating we get

$$\int_0^t f'(\tau)\, d\tau \ge 2^{p-1}\int_0^t \tau\, d\tau$$

whereby

$$f(t) \ge 2^{p-2}t^2.$$

With this in hand we can apply the Squeeze theorem to (16.1.1), which yields

$$\lim_{j\to\infty} f(B_j/A_j) = 0$$

and

$$0 \le (\frac{B_j}{A_j}) \le \frac{1}{2^{p-2}} f(B/j/A_j),$$

and so $B_j \to 0$. □

Lecture 17 Dual Spaces

17.1 Characterising Dual Spaces

Theorem 17.1.1. *Let $1 \leq p < \infty$, and let (Ω, Σ, μ) be a measure space, with the additional assumption of μ being σ-finite in the case where $p = 1$. Then $L^p(\Omega)^* \cong L^q(\Omega)$, where $1/p + 1/q = 1$. In particular for each bounded, linear functional $L \in L^p(\Omega)^*$ there exists a unique function $h \in L^q(\Omega)$ such that*

$$L(f) = \int_\Omega fg\, d\mu.$$

Proof. We treat the case of $1 < p < \infty$ first. Given a bounded/continuous linear functional $L \in L^p(\Omega)^*$, define

$$K = \{\, g \in L^(p) \mid L(g) = 0 \,\}.$$

This set K is a convex (and linear) space, and it is closed. Assume that $L \not\equiv 0$, since otherwise the result is trivial. Then there must exist some $f \in L^p(\Omega)$ such that $L(f) \neq 0$, meaning that $f \notin K$.

Therefore by the projection on a convex set, there exist some $h \in K$ such that

$$\int_\Omega k|f-h|^{p-2}(f-h)\, d\mu \leq 0$$

for every $k \in K$, where we call k a ***test function***.

Since $-k \in K$ whenever $k \in K$, since $-0 = 0$, we have

$$\int_\Omega k|f-h|^{p-2}(f-h)\, d\mu \geq 0,$$

so in all

$$\int_\Omega k|f-h|^{p-2}(f-h)\, d\mu = 0$$

for all $k \in K$. We have shown before that $|f-h|^{p-2}(f-h) \in L^q(\Omega)$.

Next, for every $g \in L^p(\Omega)$, write

$$g = \frac{L(g)}{L(f-h)}(f-h) + \tilde{g}.$$

Passing this through L, we have

$$L(g) = \frac{L(g)}{L(f-h)}L(f-h) + L(\tilde{g}) = L(g) + L(\tilde{g}),$$

implying that $L(\tilde{g}) = 0$, whereby $\tilde{g} \in K$. Note that this is somewhat an analogue to projection on orthogonal sets, except of course we have no sense of orthogonality when $p \neq 2$.

Consider

$$\begin{aligned}\int_\Omega |f-h|^{p-2}(f-h)g\, d\mu &= \int_\Omega |f-h|^{p-2}(f-h)\frac{L(g)}{L(f-h)}(f-h)\, d\mu \\ &= \frac{L(g)}{L(f-g)}\int_\Omega |f-h|^p\, d\mu = \frac{\|f-h\|_p^p}{L(f-h)}L(g).\end{aligned}$$

Therefore, for every $g \in L^p(\Omega)$, we have

$$\int_\Omega g|f-h|^{p-2}(f-h)\,d\mu = \frac{\|f-h\|_p^p}{L(f-h)}L(g).$$

Inspired by this, set

$$v = \frac{L(f-h)}{\|f-h\|_p^p}|f-h|^{p-2}(f-h)$$

which by construction is in $L^q(\Omega)$. Therefore

$$L(g) = \int_\Omega vg\,d\mu$$

for every $g \in L^p(\Omega)$.

We proceed to show that this is unique: suppose that

$$\int_\Omega (v-w)g\,d\mu = 0$$

for every $g \in L^p(\Omega)$ and some $w \in L^q(\Omega)$. Choose $g = |v-w|^{q-2}(v-w) \in L^p(\Omega)$, then

$$\int_\Omega |v-w|^q\,d\mu = 0$$

implies that $v - w = 0$ almost everywhere, which is the case if and only if $\|v-w\|_q^q = 0$.

Finally, let us show that $\|L\| = \|v\|_q$. This is straight forward:

$$\|L\| = \sup_{\|g\|_p \neq 0} \frac{|L(f)|}{\|g\|_p} = \sup_{\|g\|_p \neq 0} \frac{1}{\|g\|_p}\left|\int_\Omega vg\,d\mu\right| \leq \sup_{\|g\|_p \neq 0} \frac{1}{\|g\|_p}\|v\|_q\|g\|_p$$

with the inequality being Hölder, whereby

$$\sup_{\|g\|_p \neq 0} \frac{L(g)}{\|g\|_p} \leq \|v\|_q$$

meaning that $\|L\| \leq \|v\|_q$. For the opposite inequality, take $g = |v|^{q-2}v/\|v\|_q^{q/p}$. Therefore

$$L(g) = \int_\Omega \frac{|v|^{q-2}|v|^2}{\|v\|_q^{q/p}}\,d\mu = \frac{\|v\|_q^q}{\|v\|_q^{q/p}} = \|v\|_q^{q(1-1/p)} = \|v\|_q,$$

so $\|L\| = \|v\|_q$.

For $p = 1$, first assume that $\mu(\Omega) < \infty$. In this case a continuous linear functional $L \in L^1(\Omega)^*$ has a restriction on $L^p(\Omega)$ such that

$$|L(f)| \leq C\|f\|_1 \leq C\mu(\Omega)^{1/q}\|f\|_p$$

for all $p \geq 1$. To see this, just write

$$\int_\Omega |f|\,d\mu = \int_\Omega 1 \cdot |f|\,d\mu$$

and use Hölder's inequality.

Then by the previous case there exists a unique $v_p \in L^q(\Omega)$ such that

$$L(f) = \int_\Omega v_p f \, d\mu$$

for every $f \in L^p(\Omega)$. Since $L^r(\Omega) \subset L^p(\Omega)$ for every $r \geq p$ (again due to Hölder, when we have a finite measure space), the uniqueness of v_p for every p implies that $v = v_p$ is independent of p.

Now pick some p, q with $1/p + 1/q = 1$ and $p > 1$, and choose $f = |v|^{q-2}v$. Using Hölder's inequality on this, we have

$$\|v\|_q^q = \int_\Omega |v|^q \, d\mu = L(f) \leq C\mu(\Omega)^{1/q}\Big(\int_\Omega |v|^{(q-1)/p} \, d\mu\Big)^{1/p} = C\mu(\Omega)^{1/q}\|v\|_q^{q-1}.$$

Therefore $\|v\|_q \leq V\mu(\Omega)^{1/q}$ for every $1 < q < \infty$.

We claim that $v \in L^\infty(\Omega)$ and $\|v\|_\infty \leq C$. □

Lecture 18 Establishing Weak Convergence

18.1 Restricting to Dense Subsets

We have now established that linear bounded functionals have representation, meaning that we can represent each such functional from $L^p(\Omega)^*$ by a function from the dual space $L^q(\Omega)$, and integrating against the latter.

This means that to establish weak convergence in $L^p(\Omega)$ we would integrate against all functions in $L^q(\Omega)$, however we can do much better: it suffices to consider only some $h \in L^q(\Omega)$, namely h from a countable, dense subset, fall it $\mathcal{F}$. If we can then show that $\|f_j\|_p \leq M < \infty$, then there exists a subsequence $\{ f_{j_k} \}$ such that $f_{j_k} \rightharpoonup f \in L^p(\Omega)$.

(Compare with Bolzano-Weierstrass form real analysis, saying that if we have a bounded sequence in $\mathbb{R}^n$ or $\mathbb{C}^n$, there must exist a convergent subsequence.)

Lemma 18.1.1. *There exists a countable set of function $\mathcal{F} = \{ \varphi_1, \varphi_2, \dots \}$ such that for every $1 \leq p < \infty$ and for every measurable $\Omega \in \mathbb{R}^n$, and for every $f \in L^p(\Omega)$, we have for every $\varepsilon > 0$*

$$\|f - \varphi_j\|_p < \varepsilon$$

for some $\varphi_j \in \mathcal{F}$.

Proof. It is enough to prove it for $\Omega = \mathbb{R}^n$, since if it is true in the whole space it is true in a subspace as well.

Set

$$\Gamma_{j,m} = \{ x \in \mathbb{R}^n \mid 2^{-j} m_i < x_i < 2^{-j}(m_i + 1) \}$$

where $m_i \in \mathbb{Z}^n$ (i.e. on the integer lattice in $\mathbb{R}^n$) for $j = 1, 2, 3, \dots$. In other words, a dyadic cube on the integer lattice.

Define a preliminary class $\mathcal{F}_j$ by $f \in \mathcal{F}_j$ if and only if $f(x) = c_{j,m}$ a constant if $x \in \Gamma_{j,m}$ and $c_{j,m} \in \mathbb{Q}$.

Now let

$$\mathcal{F} = \bigcup_{j=1}^\infty \mathcal{F}_|$$

a countable union of countable sets, making it countable.

Now if $f \in L^p(\Omega)$, then there exists some continuous $\tilde{f} \in L^p(\Omega)$ such that

$$\int_\Omega |f - \tilde{f}|^p \, d\mu < \frac{\varepsilon}{3}.$$

That is to say, continuous functions are dense in $L^p(\Omega)$—we will prove this later, but for now take it as given.

We therefore need to find $f_j \in \mathcal{F}$ such that

$$\int_\Omega |\tilde{f} - f_j|^p \, d\mu < \frac{2\varepsilon}{3}.$$

Next let us assume that $\tilde{f}$ is zero outside of a fixed cube $\gamma = \{ x \,|\, -2^j \leq x_i < 2^j \}$ for $j \in \mathbb{Z}$. (If this is not the case, approximate it on a sequence of increasing nested cubes and use the Dominated convergence theorem).

Now since $\tilde{f}$ is continuous everywhere, it is particular continuous on the closure of γ, and so uniformly continuous on γ. In other words for every $\varepsilon' > 0$ there exists a $\delta > 0$ such that $|\tilde{f}(y) - \tilde{f}(x)| < \varepsilon'$ if $|x - y| < \delta$.

For every $j \in \mathbb{Z}$, define

$$\tilde{f}_j(x) = 2^{nj} \int_{\Gamma_{j,m}} \tilde{f} \, d\mu$$

if $x \in \Gamma_{j,m}$, making this a piecewise constant approximation. Now choose j such that $\sqrt{n} 2^{-j} \leq \delta$. Then

$$\int_{\mathbb{R}^n} |\tilde{f}(x) - \tilde{f}_j(x)|^p \, d\mu(x) \leq \mu(\gamma)(2\varepsilon')^p.$$

Now choose ε' small enough so that the above quantity is bounded by $\varepsilon/3$, and use this along with the triangle inequality, meaning that

$$\int_\Omega |f - \tilde{f}_j|^p \, d\mu \leq \frac{\varepsilon}{3}.$$

Finally replace $\tilde{f}_j$ with $\hat{f}_j$ that assumes only rational values, and we are done. □

Theorem 18.1.2. *Let $\Omega \subset \mathbb{R}^n$, with Ω measurable. Let $1 < p < \infty$, and suppose $\{ f_n \} \subset L^p(\Omega)$ is bounded. Then there exists a subsequence $\{ f_{n_i} \}$ and an element $f \in L^p(\Omega)$ such that $f_{n_i} \rightharpoonup f$ weakly in $L^p(\Omega)$ as $j \to \infty$.*

Proof. By our duality theorem, $L^p(\Omega)^* \cong L^q(\Omega)$ with $1/p + 1/q = 1$. We need to find f_{n_i} such that

$$\int_\Omega f_{n_i} g \, d\mu$$

converges as $i \to \infty$. By the separability lemma above, it is enough to consider only g from a countably dense family $\{ \varphi_1, \varphi_2, \dots \} \subset L^q(\Omega)$.

First consider

$$\int_\Omega f_i \varphi_q \, d\mu = c_{1,j},$$

a bounded sequence of real numbers. Therefore there exists a subsequence $\{ f_{1,j} \}$ such that

$$\int_\Omega f_{1,j}\varphi_1 \, d\mu$$

converges to c_1 as $j \to \infty$.

Now, starting with this subsequence, consider

$$c_{2,j} = \int_\Omega f_j \varphi_2 \, d\mu$$

and extract another convergent subsequence.

Continue this way, defining

$$F_k = f_{k,k}$$

for a Cantor diagonal argument. Then

$$\int_\Omega F_k \varphi_\ell \, d\mu \to c_\ell$$

as $K \to \infty$ and for all $\ell = 1, 2, 3, \ldots$.

To show existence of $f \in L^p(\Omega)$, define a linear functional

$$L(y) = \lim_{j\to\infty} \int_\Omega F_j g \, d\mu$$

for $g \in L^q(\Omega)$. This functional is bounded since $\|F_j\|_p \leq M < \infty$, so by the duality theorem there exists a unique $f \in L^p(\Omega)$ such that

$$L(g) = \int_\Omega fg \, d\mu.$$

□

Lecture 19 Important Tools in L^p Spaces

19.1 Moments

Let (Ω, Σ, μ_X) be a measure space with $\mu_X(\Omega) = 1$, making it a probability space. Let $X\colon (\Omega, \Sigma, \mu_X) \to \mathbb{R}$, called a ***random variable***.

We define the ***moments*** of this as

$$\int_\Omega x^k \, d\mu_X$$

for $k = 1, 2, 3, \ldots$.

There are some moments (or quantities related to moments) that are particularly well-known. The first moment, known as the ***expectation*** or ***mean*** of X is

$$\mathrm{E}(X) = \int_\Omega x \, d\mu_X(x).$$

The ***variance*** of X is not quite a moment, but related to the second moment:

$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{E}((X-\mathrm{E}(X))^2) = \int_\Omega (x-\mathrm{E}(X))^2\, d\mu_X(x)\\ &= \int_\Omega x^2\, d\mu_X(x) - 2\int_\Omega x\,\mathrm{E}(X)\, d\mu_X(x) + \mathrm{E}(X)^2\int_\Omega d\mu_X(x)\\ &= \int_\Omega x^2\, d\mu_X(x) - 2\,\mathrm{E}(X)^2 + \mathrm{E}(X)^2 = \int_\Omega x^2\, d\mu_X(x) - \mathrm{E}(X)^2\\ &= \mathrm{E}(X^2) - \mathrm{E}(X)^2.\end{aligned}$$

Theorem 19.1.1. *Suppose $a \le X \le b$ with probability 1. Then*

$$\mathrm{Var}(X) \le \frac{(b-a)^2}{4}.$$

Proof. First, assume $a = 0$ and $b = 1$, meaning that $0 \le X \le 1$, so $X^2 \le X$. By the above we therefore have $\mathrm{Var}(X) \le \mathrm{E}(X) - \mathrm{E}(X)^2$ since $\mathrm{E}(X^2) \le \mathrm{E}(X)$. Since $0 \le X \le 1$, and since $\int_\Omega d\mu_X = 1$, we have

$$\mathrm{E}(X) = \int_\Omega x\, d\mu_X(x) \le 1.$$

Now consider $\varphi(t) = t - t^2$ on $[0,1]$. This is of course maximised at $t = 1/2$, so

$$\mathrm{Var}(X) \le \frac{1}{4}.$$

Next let $a \le X \le b$, and consider

$$Y = \frac{X-a}{b-a}$$

on $[0,1]$. Then by the above $\mathrm{Var}(Y) \le 1/4$, and moreover

$$\mathrm{Var}(Y) = \int_\Omega \big(\frac{x-a}{b-a}\big)^2\, d\mu_X(x) - \mathrm{E}\,\big(\frac{X-a}{b-a}\big)^2.$$

The first term is

$$\frac{1}{(b-a)^2}\int_\Omega (x^2 - 2ax + a^2)\, d\mu_X(x) = \frac{1}{(b-a)^2}(\mathrm{E}(X^2) - 2a\,\mathrm{E}(X) + a^2),$$

and the second term is

$$\Big(\int_\Omega \frac{x-a}{b-a}\, d\mu_X(x)\Big)^2 = \frac{1}{(b-a)^2}(\mathrm{E}(X)-a)^2 = \frac{1}{(b-a)^2}(\mathrm{E}(X)^2 - 2a\,\mathrm{E}(X) + a^2).$$

Combining these two we have

$$\mathrm{Var}(Y) = \frac{1}{(b-a)^2}(\mathrm{E}(X^2) - \mathrm{E}(X)^2) = \frac{1}{(b-a)^2}\,\mathrm{Var}(X)$$

and so

$$\frac{1}{(b-a)^2}\,\mathrm{Var}(X) \le \frac{1}{4}$$

which means that

$$\mathrm{Var}(X) \le \frac{(b-a)^2}{4}.$$

□

19.2 Convolution

The ***convolution*** of f and g is defined as

$$f * g(x) = \int_\Omega f(x-y)g(y)\,d\mu(y).$$

We can think of this in one of two ways: either as a binary operation in the space of functions or, fixing f, as an integral operator called the convolution operator with kernel f.

Note that if μ is a translation invariant measure, then

$$\int_\Omega f(y-x)\,d\mu(y) = \int_\Omega f(y)\,d\mu(y).$$

Now if $f, g \geq 0$, and $f, g \in L^1(\mathbb{R}^n)$ (or some subset $\Omega \subset \mathbb{R}^n$), then

$$\int_\Omega f * g(x)\,d\mu = \int_\Omega \int_\Omega f(x-y)g(y)\,d\mu(y)\,d\mu(x) = \int_\Omega f(x)\,d\mu(x) \int_\Omega g(y)\,d\mu(y).$$

So we have $\|f * g\|_1 = \|f\|_1 \|g\|_1$.

This can be generalised:

Lemma 19.2.1. *Let $1 < p \leq \infty$, $f \in L^p(\mathbb{R}^n)$ and $g \in L^1(\mathbb{R}^n)$. Then $f * g \in L^p(\mathbb{R}^n)$ and $\|f * g\|_p \leq \|f\|_p \|g\|_1$.*

Proof. We may assume that $f, g \geq 0$, since $|f * g| \leq |f| * |g|$.

Now if $p = \infty$,

$$f * g(x) \leq \int_{\mathbb{R}^n} \|f\|_\infty g(x-y)\,d\mu(y) = \|f\|_\infty \int_{\mathbb{R}^n} g(x-y)\,d\mu(y) = \|f\|_\infty \|g\|_1.$$

If on the other hand $1 < p < \infty$, then we write

$$f * g(x) = \int_{\mathbb{R}^n} f(y)g(x-y)^{1/p} g(x-y)^{1/q}\,d\mu(y),$$

which by Hölder is bounded by

$$f * g(x) \leq \Big(\int_{\mathbb{R}^n} f(y)^p g(x-y)\,d\mu(y)\Big)^{1/p} \Big(\int_{\mathbb{R}^n} g(x-y)\,d\mu(y)\Big)^{1/q},$$

meaning that

$$\begin{aligned}
\|f * g\|_p &= \int_{\mathbb{R}^n} (f * g(x))^p\,d\mu(x) \\
&\leq \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} f(y)^p g(x-y)\,d\mu(y) \int_{\mathbb{R}^n} g(x-y)\,d\mu(y)\Big)^{p/q} d\mu(x) \\
&= \|f\|_p^p \|g\|_1 \|g\|_1^{p/q} = \|f\|_p^p \|g\|_1^{1+p/q},
\end{aligned}$$

so

$$\|f * g\|_p \leq \|f\|_p \|g\|_q^{(1+p/q)/p} = \|f\|_p \|g\|_1. \qquad \square$$

We can do even better, in fact:

Theorem 19.2.2 (Young's inequality)**.** *If $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$, then $f * g \in L^r(\mathbb{R}^n)$ with*

$$1 + \frac{1}{r} + \frac{1}{p} + \frac{1}{q}$$

with $1/p + 1/q \geq 1$ for $1 \leq p, q \leq \infty$ and

$$\|f * g\|_r \leq \|f\|_p \|g\|_q.$$

Note that if $r = p$, then $q = 1$ and we recover the previous lemma. Moreover q must be smaller than or equal to the dual exponent of p.

Lecture 20 Young's Inequality

20.1 Proof of Young's Inequality

Theorem 20.1.1 (Young's inequality)**.** *If $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$, then $f * g \in L^r(\mathbb{R}^n)$ with*

$$1 + \frac{1}{r} + \frac{1}{p} + \frac{1}{q}$$

with $1/p + 1/q \geq 1$ for $1 \leq p, q \leq \infty$ and

$$\|f * g\|_r \leq \|f\|_p \|g\|_q.$$

Before proving it, we note the following consequences of Hölder's inequality:

$$\int (fg)^r \, d\mu \leq \Big(\int f^{rk} \, d\mu \Big)^{1/k} \Big(\int g^{rk'} \, d\mu \Big)^{1/k'}$$

meaning that

$$\|fg\|_r \leq \|f\|_{rk} \|g\|_{rk'},$$

where if we identify $rk = p_1$ and $rk' = p_2$ we have $1/p_1 + 1/p_2 = 1/r$.

Similarly with three factors, we have

$$\begin{aligned} \int (fgh)^r \, d\mu &\leq \Big(\int f^{rk} \, d\mu \Big)^{1/k} \Big(\int (gh)^{rk'} \, d\mu \Big)^{1/k'} \\ &\leq \Big(\int f^{rk} \, d\mu \Big)^{1/k} \Big(\int g^{rk's} \, d\mu \Big)^{1/(sk')} \Big(\int h^{rk's'} \, d\mu \Big)^{1/(s'k')}. \end{aligned}$$

Thereby

$$\|fgh\|_r \leq \|f\|_{rk} \|g\|_{rk's} \|h\|_{rk's'}$$

where $rk = p_1$, $rk's = p_2$, and $rk's' = p_3$, and

$$\frac{1}{r} = \frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3}.$$

Proof. To prove Young's inequality, we write $f * g$ as

$$f * g(x) = \int f(t)^{p/r} g(x-t)^{q/r} \cdot f(t)^{p(1/p-1/r)} \cdot g(x-t)^{q(1/q-1/r)} \, dt.$$

Then we use our three factor Hölder from above:

$$f * g(x) \le \Big(\int f(t)^p g(x-t)^q \, dt\Big)^{1/r} \Big(\int f(t)^{p(1/p-1/r)r'} g(x-t)^{q(1/q-1/r)r'} \, dt\Big)^{1/r'}$$
$$\le \Big(\int f(t)^p g(x-t)^q \, dt\Big)^{1/r} \Big(\int f(t)^{p(1/p-1/r)r's} \, dt\Big)^{1/(sr')} \Big(\int g(x-t)^{q(1/q-1/r)r's'} \, dt\Big)^{1/(s'r')}.$$

Noting that the last two integrals are numbers—they don't depend on x—they come out as norms of f and g, and raising to the rth power and integrating with respect to x we get

$$\int (f*g)^r \, dx \le \iint f(t)^p g(x-t)^q \, dt \, dx \|f\|_{p(1/p-1/r)r's}^{p(1/p-1/r)} \|g\|_{q(1/q-1/r)r's'}^{q(1/q-1/r)}$$
$$= \|f\|_p^p \|g\|_q^q \|f\|_{p(1/p-1/r)r's}^{p(1/p-1/r)} \|g\|_{q(1/q-1/r)r's'}^{q(1/q-1/r)}.$$

Now we would like to choose s and s' such that these latter norms are actually p and q norms, respectively, so set

$$p\big(\frac{1}{p} - \frac{1}{r}\big)r's = p \qquad \text{and} \qquad q\big(\frac{1}{q} - \frac{1}{r}\big)r's' = q$$

so

$$\big(\frac{1}{p} - \frac{1}{r}\big)r's = 1 \qquad \text{and} \qquad \big(\frac{1}{q} - \frac{1}{r}\big)r's' = 1$$

and therefore

$$\frac{1}{s} = \big(\frac{1}{p} - \frac{1}{r}\big)r' \qquad \text{and} \qquad \frac{1}{s'} = \big(\frac{1}{q} - \frac{1}{r}\big)r'.$$

Now for this to work—i.e. for our Hölder above to be valid—we need $1/s + 1/s' = 1$, so let's try:

$$\frac{1}{s} + \frac{1}{s'} = r'\big(\frac{1}{p} - \frac{1}{r} + \frac{1}{q} - \frac{1}{r}\big) = r'\big(1 - \frac{1}{r}\big) = r'\frac{1}{r'} = 1,$$

so our s works.

Thus

$$\|f * g\|_r \le \|f\|_p^{p/r(1+1/p-1/r)} \|g\|_q^{q/r(1+1/q-1/r)}.$$

It remains to show that those exponents are 1.

This is the case since

$$p + r\big(1 - \frac{p}{r}\big) = p + r - p = r,$$

and so after dividing through by $1/r$ in the above, we have 1.

Next note that

$$\frac{1}{q} + \frac{1}{p} = \frac{1}{r'}\frac{1}{s}\frac{1}{1-p/r} + \frac{1}{r'}\frac{1}{s'}\frac{1}{1-q/r}.$$

Moreover

$$\frac{1}{r'}\frac{1}{1-p/r} = \big(1 - \frac{1}{r}\big)\frac{1}{1-p/r} = \frac{r-1}{r-p} \ge 1,$$

and similarly for q,

$$\frac{1}{q} + \frac{1}{p} \ge \frac{1}{s} + \frac{1}{s'} = 1. \qquad \square$$

Theorem 20.1.2. *If $f \in L^p(\mathbb{R}^n)$, $1 \le p < \infty$. Then*

$$\lim_{|h|\to 0} \|f(x+h) - f(x)\|_p = 0.$$

Proof. Let C_p denote the class of functions satisfying this. Clearly finite linear combinations of same are still in C_p, and moreover if $f_k \in C_p$ such that $\|f_k - f\|_p \to 0$, then $f \in C_p$.

To see this, note that

$$\|f(x+h)-f(x)\|_p \le \|f(x+h)-f_k(x+h)\|_p+\|f_k(x)-f(x)\|_p+\|f_k(x+h)-f_k(x)\|_p$$

and the first two terms in the right-hand side go to 0 since we assumed $\|f_k - f\|_p \to 0$, and the second because f_k is in C_p.

Thus

$$\limsup_{|h|\to 0} \|f(x+h) - f(x)\|_p \le 2\|f_k - f\|_p.$$

Next, the characteristic function of a cube is in C_p as well, since the difference of the cube and a shifted version of itself is 0 on their intersection and outside of both, and ± 1 in the areas they don't share, but that area is vanishingly small as $|h| \to 0$. □

Lecture 21 Approximate Identity

21.1 Approximate Identity

We define f_ε as the convolution of f against a particular kernel k_ε,

$$f_\varepsilon(x) = \int k_\varepsilon(x-t)f(t)\,dt$$

in $\mathbb{R}^n$, with

$$k_\varepsilon(x) = \varepsilon^{-n}k\big(\frac{x}{\varepsilon}\big),$$

which is thus scaled in order to maintain the volume of k for all ε. This integral operator $f \mapsto f_\varepsilon$ is an ***approximate identity***. We are interested in the behaviour of f_ε as $\varepsilon \to 0$, where $f \in L^p(\mathbb{R}^n)$ and k is from some suitable class of continuous and/or summable functions.

Theorem 21.1.1. *Let $k \in L^1(\mathbb{R}^n)$ with $\int k = 1$ and $f \in L^p(\mathbb{R}^n)$ for $1 \le p < \infty$. Then $\|f_\varepsilon - f\|_p \to 0$ as $\varepsilon \to 0$.*

This motivates why we call these approximate identities—in a strong convergence sense f_ε is indistinguishable from f, for sufficiently small ε, so it is almost the identity map.

Proof. First we remark that

$$(f_\varepsilon - f)(x) = \int k_\varepsilon(x-t)f(t)\,dt - \int k_\varepsilon(x-t)f(x)\,dt$$

since the kernel has volume 1. Thus

$$|f_\varepsilon - f|(x) \le \int |f(x-t) - f(x)||k_\varepsilon(t)|^{1/p}|k_\varepsilon(t)|^{1/p'}\,dt.$$

Applying Hölder we have

$$|f_\varepsilon - f|^p(x) \le \Big(\int |f(x-t) - f(x)|^p |k_\varepsilon(t)|\,dt\Big)\Big(\int |k_\varepsilon(t)|\,dt\Big)^{p/p'}$$
$$= \|f\|_1^{p/p'} \int |f(x-t) - f(x)|^p |k_\varepsilon(t)|\,dt.$$

Calling $\varphi(t) = \int |f(x-t) - f(x)|^p\,dx = \|f(x-t) - f(x)\|_p^p$, we then have

$$\|f_\varepsilon - f\|_p^p \le \|k\|_1^{p/p'} \int |k_\varepsilon(t)|\varphi(t)\,dt,$$

where we will call the integral I_ε for convenience.

Given $\delta > 0$, we write

$$I_\varepsilon = \int_{|t|<\delta} |k_\varepsilon(t)|\varphi(t)\,dt + \int_{|t|\ge\delta} |k_\varepsilon(t)|\varphi(t)\,dt = A_{\varepsilon,\delta} + B_{\varepsilon,\delta}.$$

Now given $\nu > 0$, choose $\delta > 0$ so small that $\varphi(t) < \nu$ if $|t| < \delta$ (which we can do due to the L^p continuity theorem from last lecture). Then

$$A_{\varepsilon,\delta} \le \nu \int_{|t|<\delta} |k_\varepsilon(t)|\,dt \le \nu\|k\|_1$$

for every $\varepsilon > 0$. Moreover φ is bounded because $|\varphi(t)| \le \|f(x-t)\|_p^p + \|f(x)\|_p^p = 2\|f\|_p^p$. This helps with estimating $B_{\varepsilon,\delta}$ since

$$B_{\varepsilon,\delta} = \int_{|t|\ge\delta} |k_\varepsilon(t)|\varphi(t)\,dt \le 2\|f\|_p^p \int_{|t|\ge\delta} |k_\varepsilon(t)|\,dt.$$

Thus

$$\int_{|t|\ge\delta} |k_\varepsilon(t)|\,dt \to 0$$

as $\varepsilon \to 0$ for every $\delta > 0$. Therefore

$$\varepsilon^{-n} \int_{|t|\ge\delta} |k(t/\varepsilon)|\,dt = \int_{|y|\ge\delta/\varepsilon} |k(y)|\,dy = \int_{\mathbb{R}^n} |k(y)|\,dy - \int_{|y|\le\delta/\varepsilon} |k(y)|\,dy$$

which goes to 0 as $\varepsilon \to 0$, whereby $B_{\varepsilon,\delta} \to 0$ as $\varepsilon \to 0$ for every $\delta > 0$.

Then $I_\varepsilon \le \nu\|k\|_1 + B_{\varepsilon,\delta}$ meaning that

$$\limsup_{\varepsilon\to 0} I_\varepsilon \le \nu\|k\|_1$$

for every $\nu > 0$, so

$$\limsup_{\varepsilon\to 0} I_\varepsilon = 0 = \lim_{\varepsilon\to 0} I_\varepsilon. \qquad \square$$

Lecture 22 Lebesgue Sets

22.1 Approximate Identities and Lebesgue Sets

Definition 22.1.1 (Lebesgue set)**.** Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be Lebesgue summable. A point $x_0 \in \mathbb{R}^n$ belongs to the ***Lebesgue set*** of f is

$$\lim_{|Q|\to 0} \frac{1}{|Q|} \int_Q |f(x_0 - x) - f(x_0)|\,dx = 0$$

for cubes Q with sides parallel to the coordinate axes and centred at 0.

Note that we can of course do the same with balls.

Theorem 22.1.2. *Let f be Lebesgue summable, k bounded and $k(x) = O(|x|^{-n-\lambda})$ with $\lambda > 0$ as $|x| \to \infty$, and moreover $\int k\,dx = 1$. If $f_\varepsilon = f * k_\varepsilon$, then $f_\varepsilon \to f$ at each point of the Lebesgue set of f.*

To prove this we will require the following lemma:

Lemma 22.1.3. *Suppose f is Lebesgue summable over a spherical shell $a \le |x| \le b$ and $\varphi(\rho)$ is continuous for $a \le \rho \le b$, for $0 \le a \le b < \infty$. Moreover let*

$$F(\rho) = \int_{a \le |x| \le \rho} f(x)\,dx$$

for $a \le \rho \le b$. Then

$$\int_{a \le |x| \le b} f(x)\varphi(|x|)\,dx = \int_a^b \varphi(\rho)\,dF(\rho).$$

Proof. Since $f = f_+ - f_-$, $F(\rho)$ is the difference of two increasing bounded functions. Thus $F(\rho)$ is of bounded variation and therefore the Riemann-Stieltjes integral in the right-hand side is well-defined. We may also assume that $f \ge 0$.

Let

$$I = \int_{a \le |x| \le b} f(x)\varphi(|x|)\,dx,$$

and partition $[a, b]$ with

$$a = \rho_0 < \rho_1 < \rho_2 < \ldots < \rho_k = b.$$

Then

$$I = \sum_{i=1}^{k} \int_{\rho_{i-1} \le |x| \le \rho_i} f(x)\varphi(|x|)\,dx.$$

In the usual fashion take $m_i = \inf \varphi(|x|)$ and $M_i = \sup \varphi(|x|)$ on the ith interval of the partition, and so

$$\sum_{i=1}^{k} m_i(F(\rho_i) - F(\rho_{i-1})) \le I \le \sum_{i=1}^{k} M_i(F(\rho_i) - F(\rho_{i-1})),$$

where these are the upper and lower Riemann-Stieltjes sums for the right-hand side integral.

These two converge under refinement of the partition, and so the lemma is proven. □

Proof of the Theorem. Fix x_0 in the Lebesgue set of f. We may assume $x_0 = 0$, otherwise we use the function $f(x_0+x)$ instead. Moreover since $k = O(|x|^{-n-\lambda})$, we have $k = o(|x|^{-n})$, and the previous theorem would apply if f was continuous at 0.

Thus by subtracting a continuous function $g(x)$ with compact support such that $g(0) = f(0)$, we may assume $f(0) = 0$.

The assumptions on k can be written as

$$|k(x)| < \frac{M_1}{(1 + |x|)^{n+\lambda}}$$

from the boundedness and the asymptotic behaviour, and so

$$|k_\varepsilon(x)| \leq M_1 \frac{\varepsilon^\lambda}{(\varepsilon + |x|)^{n+\lambda}},$$

implying that

$$|f_\varepsilon(0)| = \left| \int k_\varepsilon(x) f(x)\, dx \right| \leq M_1 \int |f(x)| \frac{\varepsilon^\lambda}{(\varepsilon + |x|)^{n+\lambda}}\, dx.$$

Now let

$$F(\rho) = \int_{|x|\leq\rho} |f(x)|\, dx.$$

Since $x_0 = 0$ is a point in the Lebesgue set and $f(0) = 0$, given $\eta > 0$ there exists $\delta > 0$ such that $F(\rho) \leq \eta\rho^n$ if $\rho \leq \delta$. We need to show that

$$\int |f(x)| \frac{\varepsilon^\lambda}{(\varepsilon + |x|)^{n+\lambda}}\, dx \to 0$$

as $\varepsilon \to 0$.

Write the integral as

$$\int_{|x|\leq\delta} |f(x)| \frac{\varepsilon^\lambda}{(\varepsilon + |x|)^{n+\lambda}}\, dx + \int_{|x|>\delta} |f(x)| \frac{\varepsilon^\lambda}{(\varepsilon + |x|)^{n+\lambda}}\, dx,$$

and call the first integral A and the second integral B.

Take

$$\varphi(\rho) = \frac{\varepsilon^\lambda}{(\varepsilon + \rho)^{n+\lambda}}$$

and let $[a, b] = [0, \delta]$. By the lemma, then,

$$A = \int_0^\delta \frac{\varepsilon^\lambda}{(\varepsilon + \rho)^{n+\lambda}}\, dF(\rho)$$

and if we integrate this by parts we get

$$A = \frac{\varepsilon^\lambda}{(\varepsilon + \delta)^{n+\lambda}} F(\delta) - \frac{\varepsilon^\lambda}{(\varepsilon + 0)^{n+\lambda}} F(0) + (n+\lambda) \int_0^\delta F(\rho) \frac{\varepsilon^\lambda}{(\varepsilon + \rho)^{n+\lambda+1}}\, d\rho.$$

The first term goes to 0 as $\varepsilon \to 0$ by just looking at the powers, the second one if 0 because $f(0) = 0$, and the last one we can evaluate as

$$\begin{aligned}(n+\lambda) \int_0^\delta F(\rho) \frac{\varepsilon^\lambda}{(\varepsilon + \rho)^{n+\lambda+1}}\, d\rho &\leq (n+\lambda)\eta \int_0^\delta \rho^n \frac{\varepsilon^\lambda}{(\varepsilon + \rho)^{n+\lambda+1}}\, d\rho \\ &= (n+\lambda)\eta \int_0^{\delta/\varepsilon} \frac{t^n}{(1+t)^{n+\lambda+1}}\, dt \leq C\eta\end{aligned}$$

with C depending only on η and λ.

Therefore

$$\limsup_{\varepsilon\to 0} A \leq C\eta.$$

To estimate B, note that if $|x| > \delta$, then $\varepsilon + |x| > \delta$, whereby

$$B \le \frac{\varepsilon^\lambda}{\delta^{n+\lambda}} \int_{|x|>\delta} |f(x)|\,dx \le \frac{\varepsilon^\lambda}{\delta^{n+\lambda}} \|f\|_1 \to 0$$

as $\varepsilon \to 0$ and δ fixed.

Thus

$$\limsup_{\varepsilon\to 0} A + B \le C\eta$$

for every $\eta > 0$, and so

$$\lim_{\varepsilon\to 0} A + B = 0. \qquad \square$$

Lecture 23 Maximal Functions

23.1 Properties of the Maximal Function

In what follows we'll let $B(x, r)$ denote the Euclidean ball of radius r centred on x, and we write $|E|$ to indicate the Lebesgue measure of the set E.

Then we can think of

$$\int_{|B(x,r)|} \int_{B(x,r)} f(y)\,dy$$

as the integral average of f around x.

Definition 23.1.1 (Hardy-Littlewood maximal function)**.** Let f be a summable function, and define

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

called the ***Hardy-Littlewood maximal function***.

Some obvious properties follow immediately from the definition. For example $M(cf) = |c|Mf$, and it is subadditive, meaning that

$$M(f_1 + f_2) \le Mf_1 + Mf_2.$$

Moreover if $f \in L^\infty(\mathbb{R}^n)$, then $Mf \le \|f\|_\infty$, so $Mf \in L^\infty(\mathbb{R}^n)$ as well.

This is not true for $f \in L^1(\mathbb{R}^n)$ in general. To see this, fix $R > 0$ and let $|x| > R$. Then

$$\begin{aligned} Mf(x) &\ge \frac{1}{|B(x, R+|x|)|} \int_{B(x,R+|x|)} |f(y)|\,dy \ge \frac{\int_{B(0,R)} |f(y)|\,dy}{|B(x, R+|x|)|} \\ &= \frac{\int_{B(0,R)} |f(y)|\,dy}{\omega_n (R+|x|)^n} = \frac{C_r}{\omega_n (R+|x|)^n} = O(|x|^{-n}) \end{aligned}$$

as $|x| \to \infty$, with ω_n being the volume of the unit ball in $\mathbb{R}^n$.

Hence this does not converge quickly enough at infinity to be in $L^1(\mathbb{R}^n)$.

Theorem 23.1.2. *If $f \in L^p(\mathbb{R}^n)$ with $1 \le p \le \infty$, we have $Mf(x) < \infty$ almost everywhere, and* (*i*) *if* $f \in L^1(\mathbb{R}^n)$, *then*

$$|\{\, x \mid Mf > t \,\}| \le \frac{5^n}{t} \int_{\mathbb{R}^n} |f| \, dt.$$

(*ii*) *If* $f \in L^p$, $1 < p \le \infty$, *then* $Mf \in L^p(\mathbb{R}^n)$ *and*

$$\|Mf\|_p \le 2 \cdot 5^{n/p} \big(\frac{p}{p-1}\big)^{1/p} \|f\|_p,$$

for $1 < p < \infty$ *and* $\|Mf\|_\infty \le \|f\|_\infty$.

To prove this we use the following covering lemma:

Lemma 23.1.3. *Let* $\mathcal{B}$ *be a family of balls in a metric space such that*

$$\sup\{\, \operatorname{diam} B \mid B \in \mathcal{B} \,\} < \infty.$$

Then there is a subfamily of pairwise disjoint balls $\mathcal{B}' subset \mathcal{B}$ *such that*

$$\bigcup_{B \in \mathcal{B}} B \subset \bigcup_{B \in \mathcal{B}'} 5B,$$

where by $5B$ *we mean the five-fold inflation of* B. *If the metric space is separable, then* $\mathcal{B}'$ *is countable.*

Proof. Let $R = \sup \operatorname{diam} B B \in \mathcal{B}$. Define

$$\mathcal{F}_j = \{\, B \in \mathcal{B} \mid \frac{R}{2^j} < \operatorname{diam} B \le \frac{R}{2^{j-1}} \,\}$$

for $j \in \mathbb{N}$. Then clearly by construction $\mathcal{B} = \bigcup_{n=1}^{\infty} \mathcal{F}_n$. Let $\mathcal{B}_1 \subset \mathcal{F}_1$ be the maximal family of pairwise disjoint balls in $\mathcal{F}_1$.

Next, suppose $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_{j-1}$ are already defined. Then define $\mathcal{B}_j$ to be the maximal family of pairwise disjoint balls in

$$\mathcal{F}_j \cap \{\, B \mid B \cap B' = \varnothing \text{ for all } B \in \bigcup_{i=1}^{j-1} \mathcal{B}_i \,\}.$$

Then define

$$\mathcal{B}' = \bigcup_{i=1}^{\infty} \mathcal{B}_i.$$

Note that every ball $B \in \mathcal{F}_j$ intersects with a ball in $\bigcup_{i=1}^{j} \mathcal{B}_i$. Suppose $B \cap B_1 \neq \varnothing$, meaning that $B_1 \in \bigcup_{i=1}^{j} \mathcal{B}_i$.

Then

$$\operatorname{diam} B \le \frac{R}{2^{j-1}} = 2\frac{R}{2^j} \le 2 \operatorname{diam} B_1,$$

implying that $B \subset 5B_1$. □

Proof of the Theorem. For (i), let $f \in L^1(\mathbb{R}^n)$ and let $E_t = \{\, x \mid Mf > t \,\}$. By the definition of Mf, for $x \in E_t$ there exists some $r_x > 0$ such that

$$\frac{1}{|B(x, r_x)|} \int_{B(x, r_x)} |f(y)| \, dy > t,$$

which rearranged yields

$$|B(x, r_x)| < \frac{1}{t} \int_{B(x,r_x)} |f(y)|\, dy.$$

Therefore

$$E_t \subset \bigcup_{x \in E_t} B(x, r_x).$$

Furthermore,

$$\sup_{x \in E_t} r_x < \infty$$

since $f \in L^1(\mathbb{R}^n)$.

By the lemma, there exist pairwise disjoint balls $B(x_i, r_{x_i})$ for $i = 1, 2, \ldots$ such that

$$E_t \subset \bigcup_{i=1}^{\infty} B(x_i, 5r_{x_i})$$

and therefore

$$|E_t| \le 5^n \sum_{i=1}^{\infty} |B(x_i, r_{x_i})| \le \frac{5^n}{t} \sum_{i=1}^{\infty} \int_{B(x_i, r_{x_i})} |f(y)|\, dy \le \frac{5^n}{t} \int_{\mathbb{R}^n} |f(y)|\, dy = \frac{5^n}{t} \|f\|_1.$$

To be continued… □

Lecture 24 More on Maximal Functions

24.1 Lebesgue Differentiation Theorem

We'll first finish off the proof from last time:

Proof continued. Let $f = f_1 + f_2$, where $f_1 = f\chi_{\{\,|f|>t/2\,\}}$ and $f_2 = f\chi_{\{\,|f|\le t/2\,\}}$. Then by Hölder's inequality $f_1 \in L^1(\mathbb{R}^n)$. Moreover $|f| < |f_1| + t/2$, and so by sublinearity of M we have $Mf \le Mf_1 + t/2$. This means that

$$\{\, x \mid Mf > t \,\} \subset \{\, x \mid Mf_1 > t/2 \,\}$$

since $Mf_1 + t/2 > t$ implies $Mf_1 > t/2$. Hence

$$|\{\, x \mid Mf > t \,\}| \le |\{\, x \mid Mf_1 > t/2 \,\}| \le \frac{2 \cdot 5^n}{t} \int |f_1|\, dx = \frac{2 \cdot 5^n}{t} \int_{\{\,|f|>t/2\,\}} |f|\, dx$$

by (i).

Now consider, using the Layer cake representation,

$$\begin{aligned} \int |Mf|^p\, dx &= p \int_0^{\infty} t^{p-1} |\{\, x \mid Mf > t \,\}|\, dt \le p \int_0^{\infty} \frac{2 \cdot 5^n}{t} \int_{\{\,|f|>t/2\,\}} |f|\, dx\, dt \\ &= 2^p 5^n \frac{p}{p-1} \int |f|^p\, dx, \end{aligned}$$

and taking pth roots we are done. □

Remark 24.1.1. If μ is a non-negative measure on $\mathbb{R}^n$, we can define

$$M\mu(x) = \sup_{r>0} \frac{\mu(B(x,r))}{|B(x,r)|}.$$

Similar to the above proof, if μ is a Borel measure, then we can show that

$$|\{\, x \mid M\mu(x) > t \,\}| \le \frac{5^n}{t} \mu(\mathbb{R}^n).$$

Theorem 24.1.2 (Lebesgue differentiation theorem)**.** *If $f \in L^1_{loc}$, meaning it's integrable on any ball, then*

$$\lim_{r\to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} f(y)\, dy = f(x)$$

almost everywhere.

Proof. By restricting to a large enough ball containing $B(x,r)$, we may assume $f \in L^1(\mathbb{R}^n)$. Denote by

$$f_r = \frac{1}{|B(x,r)|} \int_{B(x,r)} f(y)\, dy$$

and set

$$\Omega f(x) = \limsup_{r\to 0} f_r(x) - \liminf_{r\to 0} f_r(x).$$

Note that $\Omega f(x) \le M f(x)$.

Let h be a continuous function such that $\|f-h\|_1 < \varepsilon^2$, which is possible since continuous functions are dense in L^1.

Since h is continuous, $\Omega h = 0$ for all $x \in \mathbb{R}^n$, and therefore

$$\Omega f \le \Omega(f-h) + \Omega h = \Omega(f-h),$$

so

$$|\{\, x \mid \Omega f > \varepsilon \,\}| \le |\{\, x \mid \Omega(f-h) > \varepsilon \,\}| \le |\{\, x \mid 2M(f-h) > \varepsilon \,\}|$$

which, by the maximal theorem before, is bounded by

$$\frac{C}{\varepsilon} \int_{\mathbb{R}^n} |f-h|\, dx \le C\varepsilon.$$

This implies that $\Omega f(x) = 0$ for almost all x.

So $f_r(x) \to g(x)$ for almost all x as $r \to 0$, i.e. it converges to something, and we would like to show that $g = f$ almost everywhere. To do this, consider

$$\begin{aligned}
\int_{\mathbb{R}^n} |f_r - f|\, dx &= \int_{\mathbb{R}^n} \left| \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)-f(x)|\, dy \right| dx \\
&= \int_{\mathbb{R}^n} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)-f(x)|\, dy\, dx \\
&= \frac{1}{|B(0,r)|} \int_{\mathbb{R}^n} \int_{B(0,r)} |f(x+y)-f(x)|\, dy\, dx \\
&= \frac{1}{|B(0,r)|} \int_{B(0,r)} \|f_y - f\|_1\, dy
\end{aligned}$$

where $f_y(x) = f(x+y)$. Now this is bounded by

$$\frac{1}{|B(0,r)|} \sup_{y \in B(0,r)} \|f_y - f\|_1 \cdot |B(0,r)|$$

which goes to 0 as r goes to 0, since y is in the ball of radius r from the origin, forcing y to 0, and $\|f_y - f\|_1$ is continuous at $y = 0$.

Hence $f_r \to f$ in L^1 and $f_2 \to g$ almost everywhere, so $g = f$ almost everywhere. □

We can generalise this somewhat:

Theorem 24.1.3. *Let $f \in L^1_{loc}(\mathbb{R}^n)$. Then the set of non-Lebesgue points has Lebesgue measure 0.*

Lecture 25 Maximal Functions, continued

25.1 Calderón-Zygmund Decomposition

We start by proving the theorem stated at the end of last lecture.

Proof. First of all we'll define a ***representative*** of f by

$$f(x) := \limsup_{r \to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} f(y)\, dy$$

so that f is defined for all x, not just almost everywhere. Next let $c \in Q$ and fix a set E_c of $x \in \mathbb{R}^n$ for which

$$\lim_{r \to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - c|\, dy = |f(x) - c| \tag{25.1.1}$$

fails to hold, i.e. the points not belonging to the almost everywhere result of Lebesgue differentiation theorem. In other words $|E_c| = 0$, meaning that

$$E = \bigcup_{c \in \mathbb{Q}} E_c$$

also has measure 0. By definition then, (25.1.1) holds for all $x \in \mathbb{R}^n \setminus E$ and all $c \in \mathbb{Q}$. For each fixed $x \in \mathbb{R}^n \setminus E$, find a sequence $\{ c_k \} \subset \mathbb{Q}$ such that $c_k \to f(x)$ as $k \to \infty$. Passing to the limit $k \to \infty$ in (25.1.1) therefore proves the theorem. □

Theorem 25.1.1 (Calderón-Zygmund decomposition)**.** *Suppose $f \in L^1(\mathbb{R}^n)$, $f \geq 0$, and fix $\alpha > 0$. Then there exists an open set Ω and a closed set F such that*

(i) $\mathbb{R}^n = \Omega \cup F$ *and* $\Omega \cap F = \varnothing$*;*

(ii) $f \leq \alpha$ *almost everywhere on* F*;*

(iii) $\Omega = \bigcup_{k=1}^{\infty} Q_k$, *where* Q_k *are cubes with pairwise disjoint interiors and*

$$\alpha \le \frac{1}{|Q_k|} \int_{Q_k} f(x)\,dx \le 2^n \alpha$$

for $k = 1, 2, 3, \dots$.

Proof. Decompose $\mathbb{R}^n$ into the union of identical cubes $\hat{Q}_j$ so that

$$\frac{1}{|\hat{Q}_j|} \int_{Q_j} f(x)\,dx \le \alpha.$$

Now take any cube $\hat{Q}_j$ and divide it into 2^n identical cubes. Consider any cube Q' in this partition. There are two options: either

$$\frac{1}{|Q'|} \int_{Q'} f(x)\,dx > \alpha$$

or

$$\frac{1}{|Q'|} \int_{Q'} f(x)\,dx \le \alpha.$$

If the former holds, include Q' into the family Q_k. Note that

$$\alpha < \frac{1}{|Q'|} \int_{Q'} f(x)\,dx = 2^n \frac{1}{|\hat{Q}_j|} \int_{Q'} f(x)\,dx \le 2^n \frac{1}{|\hat{Q}_j|} \int_{\hat{Q}_j} f(x)\,dx \le 2^n \alpha$$

since $\hat{Q}_j$ has 2^n as large area as Q' has by construction.

If instead the latter holds, we continue subdividing Q' into 2^n identical cubes. Continue this process infinitely many times, unless terminated earlier.

Let

$$\Omega = \bigcup_{k=1}^{\infty} Q_k.$$

We need to prove $f(x) \le \alpha$ almost everywhere on $\mathbb{R}^n \setminus \Omega = F$. Since F consists of faces of cubes (which are sets of measure 0) and points x such that there exists a sequence of cubes $\tilde{Q}_i$ such that $x \in \tilde{Q}_j$, $\operatorname{diam} \tilde{Q}_i \to 0$ as $i \to \infty$ and

$$\frac{1}{|\tilde{Q}_i|} \int_{\tilde{Q}_i} f(y)\,dy \le \alpha,$$

we have by Lebesgue differentiation theorem that

$$\lim_{i\to\infty} \frac{1}{|\tilde{Q}_i|} \int_{\tilde{Q}_i} f(y)\,dy = f(x) \le \alpha$$

almost everywhere in F. □

Next, given $0 < \alpha < n$, with $n \ge 2$, define the ***Riesz potential***

$$(I_\alpha f)(x) = \frac{1}{\gamma(\alpha)} \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{n-\alpha}}\,dy$$

where

$$\gamma(\alpha) = \frac{\pi^{n/2} 2^\alpha \Gamma(\alpha/2)}{\Gamma((n-\alpha)/2)}.$$

Theorem 25.1.2. *Let $\alpha > 0$, $1 < p < \infty$, and $\alpha p < n$. Then there exists a constant $C = C_{n,p,\alpha}$ such that*

$$\|I_\alpha f\|_{p^*} \leq c\|f\|_p$$

where $p^* = np/(n - \alpha p)$.

To prove this we first prove the following lemma:

Lemma 25.1.3. *If $0 < \alpha < n$ and $\delta > 0$, then there exists a constant $C = C_{n,\alpha}$ such that*

$$\int_{B(x,\delta)} \frac{|f(y)|}{|x-y|^{n-\alpha}}\, dy \leq C\delta^\alpha Mf(x)$$

for every $c \in \mathbb{R}^n$.

Proof. For $x \in \mathbb{R}^n$ and $\delta > 0$, consider

$$A_k = B(x, \delta/2^k) \setminus B(x, \delta/2^{k+1}),$$

and compute

$$\begin{aligned}\int_{B(x,\delta)} \frac{|f(y)|}{|x-y|^{n-\alpha}}\, dy &= \sum_{k=0}^\infty \int_{A_k} \frac{|f(y)|}{|x-y|^{n-\alpha}}\, dy \leq \sum_{k=0}^\infty \left(\frac{\delta}{2^{k+1}}\right)^{\alpha-n} \int_{A_k} |f(y)|\, dy \\ &= \omega_n \sum_{k=0}^\infty \left(\frac{\delta}{2^{k+1}}\right)^{\alpha-n} \left(\frac{\delta}{2^k}\right)^n \frac{1}{|B(x,\delta/2^k)|} \int_{B(x,\delta/2^k)} |f(y)|\, dy.\end{aligned}$$

We can estimate the integral and the measure of the ball at the end by the maximal function, acquiring the upper bound

$$\omega_n Mf(x) \sum_{k=0}^\infty \frac{\delta^\alpha}{2^{2k+1}} = C\delta^\alpha Mf(x),$$

where ω_n is the unit of the n-dimensional unit ball. □

Lecture 26 Sobolev Spaces

26.1 Weak Derivative

In what follows we will require the notion of a ***multi-index***, namely an index vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ where we define $|\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_n$.

Definition 26.1.1 (Weak (partial) derivative)**.** Let $u, v \in L^1_{loc}(U)$, with $U \subset \mathbb{R}^n$ an open set, and let α be a multi-index. Then we say that v is a ***weak (partial) derivative*** of u, written $D^\alpha u = v$, if

$$\int_U u D^\alpha \varphi\, dx = (-1)^{|\alpha|} \int_U v\varphi\, dx$$

for all $\varphi \in C_c^\infty(U)$, meaning smooth functions with compact support.

Lemma 26.1.2. *If a weak derivative exists, then it is unique.*

Proof. Suppose v and $\tilde{v}$ are two different weak derivatives of u. Then

$$\int_U v\varphi\,dx = \int_U \tilde{v}\varphi\,dx$$

for all $\varphi \in C_c^\infty(U)$, meaning that $(v-\tilde{v})\varphi = 0$ almost everywhere in U. Now by choosing in particular φ such that, for each point in U it is supported in a small neighbourhood of that point, we get that $v - \tilde{v} = 0$ almost everywhere in U. □

It is maybe not obvious that such functions φ exist—this is a standard tool of importance in distribution theory, and we offer a basic construction below:

Remark 26.1.3. We construct $\varphi \in C_c^\infty(U)$ with $\varphi > 0$ on the interior of its support. Start with

$$f(t) = \begin{cases} e^{-1/t}, & t > 0 \\ 0, & t \le 0. \end{cases}$$

We claim that $f \in C^\infty(\mathbb{R})$. For that we need

$$\lim_{t \to 0^+} f^{(k)}(t) = 0.$$

To see this, note that $f'(t) = t^{-2}e^{-1/t} = 1/(t^2 e^{1/t})$, and so taking $s = 1/t$ we have

$$\lim_{t\to 0^+} f'(t) = \lim_{s\to\infty} s^2 e^{-s} = \lim_{s\to\infty} \frac{s^2}{e^s} = 0.$$

Higher order derivatives are similar—we'll always have a polynomial in the numerator and an exponential in the denominator, and so we have 0 in the limit.

Finally then take $\varphi(x) = f(1 - |x|^2)$, and then translate this as necessary.

Example 26.1.4. Take $n = 1$ and $U = (0, 2)$. Let

$$u(x) = \begin{cases} x, & 0 < x \le 1 \\ 1, & 1 < x < 2 \end{cases} \qquad \text{and} \qquad v(x) = \begin{cases} 1, & 0 < x \le 1 \\ 0, & 1 < x < 2. \end{cases}$$

Pick any $\varphi \in C_c^\infty(U)$. Using integration by parts we have

$$\begin{aligned} \int_0^2 u\varphi'\,dx &= \int_0^1 u\varphi'\,dx + \int_1^2 \varphi'\,dx \\ &= -\int_0^1 \varphi\,dx + \varphi(1) + \varphi(2) - \varphi(1) = -\int_0^2 v\varphi\,dx \end{aligned}$$

where $\varphi(2) = 0$ since it is supported only on the interior or $(0, 2)$. Hence $v = Du$, and so u is weakly differentiable. ▲

In general, using integration by parts, differentiability implies weak differentiability, and the derivatives will agree.

However not all functions are weakly differentiable: introducing a jump ruins it.

Example 26.1.5. As before, but this time take $u(x) = 2$ on $1 < x < 2$. Then u is not weakly differentiable on U, since if we suppose there is a v such that

$$\int_0^2 u\varphi'\,dx = -\int_0^2 v\varphi\,dx$$

for every $\varphi \in C_c^\infty(U)$, then

$$-\int_0^2 v\varphi\,dx = \int_0^1 \varphi'\,dx + 2\int_0^1 \varphi'\,dx = -\int_0^2 \varphi\,dx - \varphi(1).$$

Hence if we choose $\{\,\varphi_m\,\}_{m=1}^\infty$ such that $\varphi_m \in C_c^\infty(U)$, with $0 \le \varphi_m \le 1$, and $\varphi_m(1) = 1$ along with $\varphi_m(x) \to 0$ for $x \neq 1$, then by dominated convergence theorem

$$-\int_0^2 v\varphi_m\,dx = -\int_0^1 \varphi_m\,dx - 1,$$

which in the limit becomes $0 = -1$, a contradiction. ▲

26.2 Sobolev Spaces

Definition 26.2.1 (Sobolev space)**.** Let $1 \le p \le \infty$, and let $k \ge 0$ be an integer. The ***Sobolev space*** $W^{k,p}(U)$ consists of all locally summable functions $U \to \mathbb{R}$ such that for each α with $|\alpha| \le k$, $D^\alpha u$ exists in the weak sense and $D^\alpha u \in L^p(U)$.

In particular for $p = 2$ we write $W^{k,2}(U) = \mathcal{H}^k(U)$.

We also define a norm for $u \in W^{k,p}(U)$ by

$$\|u\|_{W^{k,p}(U)} = \begin{cases} \sum_{|\alpha|\le k} \|D^\alpha u\|_p, & 1 \le p < \infty \\ \sum_{|\alpha|\le k} \operatorname{ess\,supp}|D^\alpha u|, & p = \infty. \end{cases}$$

One large reason for why we care about these weak derivatives is that under this norm, this is a normed space, as opposed to working with pointwise differentiable functions.

Example 26.2.2. Consider $U = B(0,1) \subset \mathbb{R}^n$, and $u(x) = |x|^{-\alpha}$ for $x \in U \setminus \{\,0\,\}$ and $\alpha > 0$.

We ask ourselves for what α, n, and p we have $u \in W^{1,p}(U)$. For $x \neq 0$, we have

$$\partial_{x_i} u(x) = \frac{-\alpha x_i}{|x|^{\alpha+2}},$$

and moreover $|Du(x)| = |\alpha|/|x|^{\alpha+1}$. Let $\varphi \in C_c^\infty(U)$ and fix $\varepsilon > 0$. Then

$$\int_{U\setminus B(0,\varepsilon)} u\partial_{x_i}\varphi\,dx = -\int_{U\setminus B(0,\varepsilon)} \varphi\partial_{x_i} u\,dx + \int_{\partial B(0,\varepsilon)} u\varphi\eta^i\,ds$$

where η^i is the unit normal inward. If $\alpha + 1 < n$, then $|Du(x)| \in L^1(U)$ by switching to polar coordinates and evaluating.

Then

$$\left|\int_{\partial B(0,\varepsilon)} u\varphi\eta^i\,ds\right| \le \|\varphi\|_\infty \int_{\partial B(0,\varepsilon)} \varepsilon^{-\alpha}\,ds \le C\varepsilon^{n-1-\alpha} \to 0$$

as $\varepsilon \to 0$ if $\alpha < n - 1$, since u is radial.

Hence $Du(x) \in L^p$ if $(\alpha + 1)p < n$. ▲

Lecture 27 Approximations in Sobolev Spaces

27.1 Modes of Convergence

Theorem 27.1.1. *Let*

$$\eta(x) = \begin{cases} C\exp(\frac{1}{|x|^2-1}), & \text{if } |x| < 1, \\ 0 & \text{if } |x| \geq 1, \end{cases}$$

with C chosen so that $\int \eta\, dx = 1$. Let $\eta_\varepsilon = \varepsilon^{-n}\eta(x/\varepsilon)$ and $f_\varepsilon = \eta_\varepsilon * f$. Then

(*i*) *For every $\varepsilon > 0$, $f_\varepsilon \in C^\infty(U_\varepsilon)$, where $U_\varepsilon = \{ x \in U \mid d(x, \partial U) > \varepsilon \}$, i.e. U deflated a bit.*

(*ii*) *If $f \in C(U)$, then $f_\varepsilon \to f$ uniformly on compact subsets of U.*

(*iii*) *If $f \in L^p_{loc}(U)$, with $1 \leq p < \infty$, then $f_\varepsilon \to f$ in $L^p_{loc}(U)$.*

(*iv*) *For almost all Lebesgue points x of f, $f_\varepsilon \to f$.*

(*v*) *If $f \in W^{1,p}_{loc}(U)$, with $1 \leq p \leq \infty$, then*

$$\frac{\partial f_\varepsilon}{\partial x_i} = \eta_\varepsilon * \frac{\partial f}{\partial x_i}$$

on U_ε, meaning that $f_\varepsilon \to f$ in $W^{1,p}_{loc}(U)$.

Theorem 27.1.2 (Local approximation)**.** *Let $f \in W^{1,p}(U)$, $1 \leq p < \infty$, then there exists $\{ f_k \}_{k=1}^\infty \subset W^{1,p}(U) \cap C^\infty(U)$ such that $f_k \to f$ in $W^{1,p}(U)$.*

Proof. Fix $\varepsilon > 0$. Define $U_0 = \varnothing$, and for $k = 1, 2, 3, \ldots$, define

$$U_k = \{ x \in U \mid d(x, \partial U) > \frac{1}{k} \} \cap B(0, k).$$

Set $V_k = U_{k+1} \setminus \overline{U_k}$. Let $\{ \zeta_k \}_{k=1}^\infty$ be such that

(*i*) $\zeta_k \in C_c^\infty(V_k)$;

(*ii*) $0 \leq \zeta_k \leq 1$, for $k = 1, 2, 3, \ldots$;

(*iii*) $\sum_{k=1}^\infty \zeta_k = 1$.

These three conditions make $\{ \zeta_k \}$ a so-called ***partition of unity***, the existence of which we won't prove there. Note also that the sum in (*iii*) isn't an infinite sum; the support of ζ_k overlap only for finitely many k on a given x.

It is a fact that $f\zeta_k \in W^{1,p}(U)$ with support inside V_k, implying that there exists $\varepsilon_k > 0$ such that $\operatorname{supp}(\eta_{\varepsilon_k} * (f\zeta_k)) \subset U_k$.

Now define the approximand

$$f_\varepsilon = \sum_{k=1}^\infty \eta_{\varepsilon_k} * (f\zeta_k),$$

which like in (*iii*) is a locally finite sum, so we have no issue with convergence. Now

$$\|f_\varepsilon - f\|_{L^p(U)} \leq \sum_{k=1}^\infty \Big(\int_U |\eta_{\varepsilon_k} * (f\zeta_k) - f\zeta_k|^p\, dx \Big)^{1/p},$$

where we can interchange summation and integration again because the sum is locally finite. Similarly

$$\|Ff_\varepsilon - Df\|_{L^p(U)} \le \sum_{k=1}^{\infty} \Big(\int_U |\eta_{\varepsilon_k} * D(f\zeta_k) - D(f\zeta_k)|^p \, dx \Big)^{1/p}.$$

Note that ε_k can be chosen such that

$$\Big(\int_U |\eta_{\varepsilon_k} * (f\zeta_k) - f\zeta_k|^p \, dx \Big)^{1/p} < \frac{\varepsilon}{2^k}$$

and similarly for $D(f\zeta_k)$.

Thus

$$\|f_\varepsilon - f\|_{L^p(U)} \le \varepsilon \sum_{k=1}^{\infty} \frac{1}{2^k} = \varepsilon$$

and likewise for $\|Df_\varepsilon - Df\|_{L^p(U)}$. Hence $f_\varepsilon \to f$ and $Df_\varepsilon \to Df$, both in $L^p(U)$. □

Theorem 27.1.3 (Global approximation)**.** *Suppose $U \subset \mathbb{R}^n$ is bounded, ∂U is Lipschitz (meaning that it can be locally approximated by Lipschitz functions). Then if $f \in W^{1,p}(U)$ for $1 \le p < \infty$, there exists $\{ f_k \} \subset W^{1,p}(Y) \cap C^\infty(\overline{U})$ such that $f_k \to f$ in $W^{1,p}(U)$.*

Proof. Let $x \in \partial U$ and $r > 0$. Let $\gamma \colon \mathbb{R}^{n-1} \to \mathbb{R}$ give the equation of ∂U near x, which then by assumption is Lipschitz. This then means that it is differentiable Lebesgue almost everywhere, which in turn means that we have a well-defined gradient on the boundary.

Let $Q(x, r) = \{ y \mid |y_i - x_i| < r, i = 1, 2, \dots \}$, with r small enough so that γ "works".

Write $Q = Q(x, r)$, and $Q' = Q(x, r/2)$. Suppose first that $f = 0$ near $\partial Q' \cap U$. For $y \in U \cap Q'$, $\alpha > 0$, define $y_\varepsilon = y + \varepsilon\alpha e_n$, where e_n is the Euclidean unit vector in the local coordinates adjusted to ∂U, so e_n is parallel to the normal of ∂U at x.

Note that $B(y_\varepsilon, \varepsilon) \subset U \cap Q$ for every ε small enough. Define

$$f_\varepsilon(y) = \frac{1}{\varepsilon^n} \int_U \eta(z/\varepsilon) f(y^\varepsilon - z) \, dz = \frac{1}{\varepsilon^n} \int_{B(y_\varepsilon, \varepsilon)} \eta\big(\frac{y - w}{\varepsilon} + \alpha e_n\big) f(w) \, dw$$

for $y \in U \cap Q'$.

It is a fact that $f_\varepsilon \in C^\infty(\overline{U \cap Q})$, $f_\varepsilon \to f$ in $W^{1,p}(U \cap Q')$ since $f = 0$ near $\partial U \cap Q'$ for ε small enough. Extend f_ε by 0 to $U \setminus Q'$. □

Lecture 28 GNS Inequality

28.1 Sobolev Conjugates

Suppose we are in $\mathbb{R}^n$, with a function in $W^{1,p}(\mathbb{R}^n)$, where $1 \le p < n$.

We are interested in obtaining an equality of the form

$$\|u\|_{L^q(\mathbb{R}^n)} \le C \|Du\|_{L^p(\mathbb{R}^n)}, \tag{28.1.1}$$

with q as yet unspecified—we'll get to it in due course.

Let $u_\lambda(x) = u(\lambda x)$, i.e. a scaled version of u. Then

$$\|u_\lambda\|_q \le C\|Du_\lambda\|_p.$$

Then (28.1.1) implies

$$\frac{1}{\lambda^{n/q}}\|u\|_q \le \frac{\lambda}{\lambda^{n/p}} C\|Du\|_p$$

which in turn implies

$$\|u\|_q \le \lambda^{1-n/p+n/q} C\|Du\|_p,$$

meaning that (28.1.1) can work only if $1 - n/p + n/q = 0$, meaning that $q = np/(n-p)$.

Definition 28.1.1 (Sobolev conjugate)**.** Given p, we define $q = p^* = np/(n-p)$, called the ***Sobolev conjugate*** of p.

Note that $q > p$.

Theorem 28.1.2. *There exists a* C *depending only on* p *and* n *such that* (28.1.1) *holds with* $q = p^*$.

Proof. By the approximation theorems from last time it suffices to consider $f \in C_c^\infty(\mathbb{R}^n)$.

Now by the fundamental theorem of calculus in each variable separately we have

$$f(x_1, \dots, x_i, \dots, x_n) = \int_{-\infty}^{x_i} \partial_{x_i} f(x_1, \dots, t_i, \dots, x_n)\, dt_i.$$

Now

$$|f(x)| \le \int_{-\infty}^{\infty} |Df(x_1, \dots, t_i, \dots, x_n)|\, dt_i$$

for all $i = 1, 2, \dots, n$, and so moreover

$$|f(x)^{n/(n-1)}| \le \prod_{i=1}^{n} \Big(\int_{-\infty}^{\infty} |Df(x_1, \dots, t_i, \dots, x_n)|\, dt_i \Big)^{1/(n-1)}.$$

We now integrate this in x_1, yielding

$$\int_{-\infty}^{\infty} |f|^{n/(n-1)}\, dx_1 \le \Big(\int_{\mathbb{R}} |Df|\, dt_1 \Big)^{1/(n-1)} \int_{\mathbb{R}} \prod_{i=2}^{n} \Big(\int_{\mathbb{R}} |Df|\, dt_i \Big)^{1/(n-1)} dx_1$$

which using Hölder on the second term in the right-hand side gives us

$$\le \Big(\int_{\mathbb{R}} |Df|\, dt_1 \Big)^{1/(n-1)} \Big(\prod_{i=2}^{n} \int_{\mathbb{R}} \int_{\mathbb{R}} |Df|\, dx_1\, dt_i \Big)^{1/(n-1)}$$

Repeating this for x_2 we get

$$\iint_{\mathbb{R}} |f|^{n/(n-1)}\, dx_1\, dx_2 \le \Big(\iint_{\mathbb{R}} |Df|\, dx_1\, dx_2 \Big)^{1/(n-1)} \Big(\iint_{\mathbb{R}} |Df|\, dx_2\, dt_1 \Big)^{1/(n-1)} \cdot$$
$$\cdot \Big(\prod_{i=3}^{n} \iiint_{\mathbb{R}} |Df|\, dx_1\, dx_2\, dt_i \Big)^{1/(n-1)}.$$

We keep this going, eventually getting

$$\int_{\mathbb{R}^n} |f|^{n/(n-1)}\,dx \le \Big(\int_{\mathbb{R}^n} |Df|\,dx\Big)^{n/(n-1)}.$$

Now note that if $p = 1$, then $p^* = n/(n-1)$, so the above inequality is equivalent with $\|f\|_{1^*}^{1^*} \le \|Df\|_1^{1^*}$ meaning that $\|f\|_{1^*} \le \|Df\|_1$. Therefore (28.1.1) holds for $p = 1$, for which $C = 1$.

Next, if $1 < p < n$, consider $g = |f|^\gamma$, with γ to be specified. Apply the $p = 1$ inequality fo g, so that

$$\begin{aligned}\Big(\int_{\mathbb{R}^n} |f|^{\gamma n/(n-1)}\,dx\Big)^{(n-1)/n} &\le \gamma \int_{\mathbb{R}^n} |f|^\gamma |Df|\,dx \qquad (28.1.2)\\ &\le \gamma\Big(\int_{\mathbb{R}^n} |f|^{(\gamma-1)/(p-1)p}\,dx\Big)^{(p-1)/p}\Big(\int_{\mathbb{R}^n} |Df|^p\,dx\Big)^{1/p}.\end{aligned}$$

We now with to choose γ such that $\gamma n/(n-1) = (\gamma-1)/(p-1)p$:

$$\frac{\gamma}{\gamma-1} = \frac{p(n-1)}{n(p-1)},$$

meaning that

$$\frac{\gamma-1}{\gamma} = 1 - \frac{1}{\gamma} = \frac{n(p-1)}{p(n-1)}$$

so in turn

$$\gamma = \frac{p(n-1)}{n-p},$$

which means that

$$\frac{\gamma n}{n-1} = p^* = \frac{pn}{n-p}.$$

Then (28.1.2) becomes

$$\Big(\int_{\mathbb{R}^n} |f|^{p^*}\,dx\Big)^{(n-1)/n} \le \frac{p(n-1)}{n-p}\Big(\int_{\mathbb{R}^n} |f|^{p^*}\,dx\Big)^{(p-1)/p}\Big(\int_{\mathbb{R}^n} |Df|^p\,dx\Big)^{1/p}$$

whence

$$\Big(\int_{\mathbb{R}^n} |f|^{p^*}\,dx\Big)^{(n-1)/n-(p-1)/p} \le \frac{p(n-1)}{n-p}\|Df\|_p.$$

The exponent on the left-hand integral is in fact $1/p^*$, so this is

$$\|f\|_{p^*} \le \frac{p(n-1)}{n-p}\|Df\|_p.$$

Hence the constant C is

$$C = \frac{p(n-1)}{n-p}$$

and in particular note that this goes to infinity as p goes to n. □

This is known as the ***Gagliardo–Nirenberg–Sobolev inequality***.

If instead of the whole of $\mathbb{R}^n$ we consider some open and bounded subset $U \subset \mathbb{R}^n$, where, say, $\partial U \in C^1$, we can do better:

$$\|u\|_{L^p(U)} \le C\|u\|_{W^{1,p}(U)}$$

where C depends on U, p, and n.

Proof. Having $\partial U \in C^1$ and, say, $u \in C^\infty(U) \cap W^{1,p}(U)$ means that there exists an extension Eu of u such that $Eu = u$ on U and $Eu \in C_c^\infty(\mathbb{R}^n)$ and

$$\|Eu\|_{W^{1,p}(\mathbb{R}^n)} \leq C_0 \|u\|_{W^{1,p}(U)}.$$

Next there exists $u_m \in C_c^\infty(\mathbb{R}^n)$ such that $u_m \to Eu$ in $W^{1,p}(\mathbb{R}^n)$. The GNS inequality now implies that

$$\|u_m - u_\ell\|_{L^{p^*}(\mathbb{R}^n)} \leq C_1 \|Du_m - Du_\ell\|_{L^p(\mathbb{R}^n)},$$

i.e. $u_m \to Eu$ in $L^{p^*}(\mathbb{R}^n)$.

Also $\|u_m\|_{p^*} \leq C_1 \|Du_m\|_p$ implies that

$$\|Eu\|_{L^{p^*}(\mathbb{R}^n)} \leq C_1 \|DEu\|_{L^p(\mathbb{R}^n)}$$

and $\|u\|_{L^{p^*}(U)} \leq \|Eu\|_{L^{p^*}(\mathbb{R}^n)}$. Hence

$$\|Eu\|_{L^{p^*}(\mathbb{R}^n)} \leq C_0 C_1 \|u\|_{W^{1,p}(U)}.$$

Also, if $u \in W_0^{1,p}(U)$, which in addition is the closure of the set of functions compactly supported in U with respect to the Sobolev norm, which in turn means that they have zero trace on the boundary of U, we have

$$\|u\|_{L^q(U)} \leq C(q,p,n) \|Du\|_{L^p(U)}$$

where $1 \leq q \leq p^*$. □

Index