Real Analysis Midterm Study Guide

Eduardo Corona

October 30, 2008

Contents

1 Building the Lebesgue Measure 1

2 Measurable Functions 4
  2.1 Definition and Properties 4
  2.2 Littlewood’s Principles 4

3 The Lebesgue Integral 5
  3.1 Quick Recall: The Riemann Integral 5
  3.2 Definition and Construction of the Lebesgue Integral 6
  3.3 Convergence in Measure 9

4 Differentiation and Integration 10
  4.1 Second FTC: Wiener Covers, H-L Maximal Function, Lebesgue Set 10
  4.2 First FTC: AC, BV and Vitali Cover 12
  4.3 Convex Functions and Jensen’s Lemma 16

5 $L^p$ spaces 17

1 Building the Lebesgue Measure

In the early 1900’s, Henri Lebesgue came up with a new theory of integration that was more general than Riemann integration and avoided most of its pitfalls. In order to do so, he had to develop a notion of what a measurable set is: the objective was to define this for a large class of subsets of $\mathbb{R}$, if not for all of them. This is the same as defining a function $\mu : M \to \mathbb{R}^+$, where $M \subset P(\mathbb{R})$, such that:

1. $\mu(I) = |I|$ for all $I$ intervals (so $\mu([a,b]) = b - a$)
2. $\mu(E) = \mu(E + \{x\}) \forall x \in \mathbb{R}, E \in M$ (translation invariant)
3. For any countable, pairwise disjoint \( \{E_n\}_{n=1}^{\infty} \subset M \), we want
\[
\mu(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu(E_n)
\]

**Remark 1** It is impossible to find such a function if we restrict it to \( \mathbb{Q} \).

**Remark 2** We will also see it is impossible to find such a function if we impose \( M = P(\mathbb{R}) \). This is called the "hard problem of measure in \( \mathbb{R} \)." However, it is possible to define it for a substantially big family \( M \) of "measurable sets".

In particular, we want \( M \) to be a \( \sigma \)-Algebra, that is, a family closed under countable unions and complements, which if non-empty contains both the empty set and the entire space. That is:

**Definition 3** Let \( X \) be any set, \( A \subset P(X) \). We say \( A \) is an Algebra of sets of \( X \) if:

(i) \( B, C \in A \implies B \cup C \in A \)
(ii) \( B \in A \iff B^c \in A \)

The first condition makes \( A \) a ring, and if it holds for countable unions, a \( \sigma \)-ring. Finally, if \( A \) is a \( \sigma \)-ring and (ii) holds, \( A \) is a \( \sigma \)-Algebra.

It follows from these definitions that a \( \sigma \)-Algebra is closed under countable intersections and symmetric differences as well. Other properties include:

1. **(Monotonicity)** If \( B \subset C \) then \( \mu(B) \leq \mu(C) \)

2. **(Countable Subadditivity)** For any countable collection \( \{E_n\}_{n=1}^{\infty} \subset M \),
\[
\mu(\bigcup_{n=1}^{\infty} E_n) \leq \sum_{n=1}^{\infty} \mu(E_n)
\]

3. For any countable collection \( \{E_n\}_{n=1}^{\infty} \subset M \), there exists \( \{B_n\}_{n=1}^{\infty} \subset M \) pairwise disjoint such that \( \bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} B_n \). \( (B_n = E_n \setminus \bigcup_{n=1}^{n-1} B_n) \)

We want \( M \) to be an \( \sigma \)-Algebra that at least contains all intervals in \( \mathbb{R} \). The smallest such \( \sigma \)-Algebra, or the \( \sigma \)-Algebra generated by these sets is called \( B \) the Borel \( \sigma \)-Algebra. The \( M \) that we are looking for most therefore contain \( B \) as a subset.

**Definition 4** Let \( Y \) be a subset of \( X \). Then the smallest \( \sigma \)-Algebra that contains \( Y \) is denoted by \( S(Y) \), or the \( \sigma \)-Algebra generated by \( Y \). Then,
\[
S(Y) = \bigcap\{S \mid Y \subset S \text{ and } S \text{ is a } \sigma \text{-Algebra}\}
\]

Now, to construct our set \( M \), we are going to define something called an "outer measure", which is defined for all subsets of \( \mathbb{R} \). Then, we define \( M \) by using an extension property devised by Constantin Caratheodory.
Definition 5 (Outer Measure for $\mathbb{R}$) We define $\mu^*: P(\mathbb{R}) \to \mathbb{R}^+$ as:

$$\mu^*(A) = \inf \{ \sum_{n=1}^{\infty} |I_n| \mid A \subset \bigcup_{n=1}^{\infty} I_n, \text{ open intervals} \}$$

Some properties of $\mu^*$ include:

1. $\mu^*$ is translation invariant
2. $\mu^*(\emptyset) = 0$
3. $\mu^*(\{x\}) = 0 \forall x \in \mathbb{R}$ (singletons)
4. $\mu^*(C) = 0$ for any countable set $C$
5. $A \subset B \implies \mu^*(A) \leq \mu^*(B)$ (monotone)
6. $\mu^*(I) = |I|$ for all intervals
7. $\mu^*$ is countably sub-additive

Proposition 6 $\forall A \in P(\mathbb{R}), \forall \varepsilon > 0 \exists U \in \tau_{\mathbb{R}}$ such that $A \subset U$ and $\mu^*(U) \leq \mu^*(A) + \varepsilon$.

If $\mu^*$ was countably additive, we would be done. However, this is not so, and we can show this by constructing a set (the Vitali set) such that it can't be "measured".

Example 7 We partition the interval $[0, 1)$ into a countable number of mutually exclusive $E_n$'s, each such that it must have the same measure. We fix $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, and we define $x + \alpha \mod 1$ as the fractional part of this number $\forall n \in \mathbb{Z}$. Then, we define the equivalence relation: $x \sim y \iff x - y \in \alpha \mathbb{Z} \mod(1)$. This induces a partition in equivalence classes, and by the axiom of choice, we may pick one representative $x_\lambda$ from each class $E_\lambda$. We define $V = \{x_\lambda \} \forall \lambda \in \Lambda$ to be our "problematic" set. The problem with $V$ is that, if we now define $V_n = V + \alpha \mod(1)$, $\mu^*(V_n) = \mu^*(V)$ $\forall n$ by translation invariance. If $\mu^*$ was countably additive, this would either imply everything has measure 0, or everything has measure $\infty$.

To achieve all properties we want on a measure, we restrict $\mu^*$ to the following set:

$$M = \{ E \in P(\mathbb{R}) \mid \mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c) \forall A \subset P(\mathbb{R}) \}$$

This is known as the Caratheodory Criterion, and we find that $M$ is a $\sigma$-Algebra that contains $B$, and more importantly, that $\mu^* |_M$ is a measure. Therefore, from now on, we call this restriction the Lebesgue Measure $\mu$ (or $\lambda$), and we call $M$ the Lebesgue measurable sets.
2 Measurable Functions

2.1 Definition and Properties

Definition 8. A function $f$ is Lebesgue Measurable if $\forall y \in \mathbb{R}, f^{-1}((y, \infty)) \in M$. This condition is equivalent to asking $f^{-1}((y, \infty)), f^{-1}([-\infty, y))$ or $f^{-1}([-\infty, y]) \in M$.

Facts:

1. If $f$ is measurable, then $f^{-1}(I) \in M$. This in turn implies $f^{-1}(E) \in M \forall E \in B$. However, it is not always true that $f^{-1}(E) \in M \forall E \in M$.

2. Continuous and hemicontinuous functions are measurable.

3. Step functions are measurable

4. Simple functions, that is, $f = \sum_{i=1}^{n} a_i \chi_{A_i}(x), A_i \in M$, are measurable.

5. $f, g$ measurable $\implies af + bg$ is measurable.

6. $f, g$ measurable $\implies fg$ is measurable.

7. $\{f_n\}_{n \in I}$ measurable $\implies \sup_{i \in I} \{f_i(x)\}, \inf_{i \in I} \{f_i(x)\}, \lim_{n \to \infty} \{f_n\}$ are measurable.

8. Finally, by (7), the pointwise limit of measurable functions is measurable.

To develop the Lebesgue integral, it is important to distinguish the case in which a certain property holds $\forall x \in \mathbb{R}$ except for a set of null measure. In this case, most results that require this property will still hold, since the fact that this "bad" set is null often makes it negligible:

Definition 9. We say that the property $P(x)$ (e.g. $f(x) = g(x)$) holds Almost Everywhere (AE) if $\mu \{x : P(x) \text{ is false} \} = 0$.

2.2 Littlewood’s Principles

1. Every measurable set of finite measure is approximately a finite union of intervals. That is: $\forall E \subset M \ s.t. \ \mu(E) < \infty, \ \forall \varepsilon > 0 \ \exists \{I_n\}_{n=1}^{N}$ open intervals s.t.:

   $\mu \left( \bigcup_{n=1}^{N} I_n \triangle E \right) < \varepsilon$

2. (Lusin Theorem) Every measurable function on $[a, b]$ is approximately continuous. That is: $\forall \varepsilon > 0, \eta > 0 \ \exists g \in C([a, b]) \ s.t.$

   $\mu \{ x : |f(x) - g(x)| > \varepsilon \} < \eta$
3. (Egorov’s Theorem) Every pointwise A.E. convergent sequence of measurable functions is "nearly" uniformly convergent: That is, \( f_n : [a, b] \to \mathbb{R} \) sequence of measurable functions which converge A.E. to \( f \). Then \( \forall \varepsilon > 0, \exists A \subset [a, b] \) s.t. \( \mu([a, b] \setminus A) < \varepsilon \) and \( f_n \to f \) uniformly on \( A \).

**Proof.** (1, Sketch) Given \( E \in M, \mu(E) < \infty \), given \( \varepsilon > 0 \) we know \( \exists I_n^\varepsilon \) of open, disjoint intervals such that \( \sum_{n=1}^{\infty} \frac{|I_n^\varepsilon|}{k} < \mu(E) + \varepsilon/2 \).

Then, \( O = \bigcup I_n^\varepsilon \) does the trick. ■

**Proof.** (2, Sketch) \( 0 = \mu(\bigcap_{M=1}^{\infty} \{ |f| \geq M \}) = \lim_{M \to \infty} \mu(\{ |f| \geq M \}) \). Thus, \( \exists M_\varepsilon \) such that \( E = \mu(\{ |f| \geq M_\varepsilon \}) < \varepsilon \).

On \( [a, b] \setminus E \), we approximate by simple, then step and then finally continuous functions through careful surgery. (look for proof in Grabs notes, too). ■

**Proof.** (3, Sketch) We rewrite the set where pointwise convergence criterion holds as:

\[
E = \{ x : f_n(x) \to f(x) \} = \bigcap_{k=1}^{\infty} \bigcap_{N=1}^{\infty} \bigcap_{n \geq N} \{ x : |f_n(x) - f(x)| < \frac{1}{k} \}
\]

\[
\mu(E^c) = \mu(\bigcup_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcup_{n \geq N} \{ x : |f_n(x) - f(x)| < \frac{1}{k} \}) = 0
\]

\[
\Rightarrow \mu(\bigcap_{N=1}^{\infty} \bigcup_{n \geq N} \{ x : |f_n(x) - f(x)| < \frac{1}{k} \}) = 0 \quad \forall k
\]

\[
\Rightarrow \lim_{N \to \infty} \mu(\bigcup_{n \geq N} \{ x : |f_n(x) - f(x)| < \frac{1}{k} \}) = 0 \quad \forall k
\]

If we call \( A_{N,k} = \bigcup_{n \geq N} \{ x : |f_n(x) - f(x)| < \frac{1}{k} \} \), then we find \( A_{N,k} \) s.t. \( \mu(A_{N,k}) \leq \eta 2^{-k} \), and \( A = \bigcup A_{N,k} \) does the trick. ■

### 3 The Lebesgue Integral

3.1 Quick Recall: The Riemann Integral

Let \( f : [a, b] \) bounded, \( P = \{ a = t_0 \leq t_1 \leq \ldots \leq t_n = b \} \) partition of \( [a, b] \), \( I_i = [t_{i-1}, t_i] \). We now define the following step functions (Lower and Upper):

\[
\varphi^I_L = \sum_{i=1}^{n} \inf_{x \in I_i} \{ f(x) \} \chi_{I_i}(x)
\]

\[
\varphi^I_U = \sum_{i=1}^{n} \sup_{x \in I_i} \{ f(x) \} \chi_{I_i}(x)
\]

Now, the integral of these step functions is well defined, and we identify them as sums from below and from above. We then take the limit as the partition
gets arbitrarily refined:

\[
\int \varphi^\Pi(x)dx = \sum_{i=1}^{n} \left[ \inf_{\varphi \in \Pi} \{ f(x) \} \right][I_i] = s(f, \Pi) \rightarrow R \int_a^b f(x)dx = \sup_{\Pi} \{ s(f, \Pi) \}
\]

\[
\int \varphi^\Omega(x)dx = \sum_{i=1}^{n} \left[ \sup_{\varphi \in \Omega} \{ f(x) \} \right][I_i] = S(f, \Omega) \rightarrow R \int_a^b f(x)dx = \inf_{\Omega} \{ S(f, \Omega) \}
\]

Finally, we say \( f \) is Riemann Integrable, or \( f \in R([a,b]) \) if \( R = \overline{R} \).

**Remark 10** This characterization of the lower and upper Riemann Integrals is equivalent to:

\[
\sup_{\Pi} \{ s(f, \Pi) \} = \sup_{\varphi \leq f, \ \varphi \ \text{step}} \int_a^b \varphi(x)dx
\]

\[
\inf_{\Omega} \{ S(f, \Omega) \} = \inf_{\varphi \geq f, \ \varphi \ \text{step}} \int_a^b \varphi(x)dx
\]

The only difference with the Lebesgue Integral, which will be a substantially significant one, is that we will take \( \varphi \) in the larger, more general family of simple functions.

### 3.2 Definition and Construction of the Lebesgue Integral

1. We first want to define the integral for functions such that its formulation is straightforward, that is, for simple functions. We start with:

\[
\int \chi_A d\mu = \mu(A) \text{ for any } A \in M
\]

2. For a simple function \( f = \sum_{i=1}^{n} a_i \chi_{A_i}(x) \), we would like \( \int f d\mu = \sum_{i=1}^{n} a_i \mu(A_i) \). This is true, but we go at pains to prove it is well defined. We prove it for:

- **Canonical Form**: distinct \( \{ a_i \} \), mutually exclusive \( A_i \)

- **Standard Form**: to the canonical form, if \( A = \bigcup A_i \), we add \( 0 \chi_{A^c}(x) \).

Then, we prove \( \int f + g = \int f + \int g \). Finally, we show that this works for any simple function, not just for the canonical form.

3. For **bounded functions on sets of finite measure**: Let \( f \) be defined on a set \( E \) s.t. \( \mu(E) < \infty \), \( f \) bounded. Then:

\[
\sup_{\varphi \leq f, \varphi \ \text{simple}} \left\{ \int_E \varphi(x)dx \right\} = \inf_{\psi \geq f, \psi \ \text{simple}} \left\{ \int_E \psi(x)dx \right\} \iff f \text{ is measurable}
\]

**Proof.** \((\Longleftarrow) f \text{ measurable, } N \text{ arbitrary, } |f| \leq M. \text{ We define } A_k = f^{-1}(\left[ \frac{k}{N}, \frac{k+1}{N} \right)), \ varphi_N = \sum_{k} \frac{k}{N} \chi_{A_k} \leq f \leq \sum_{k} \frac{k+1}{N} \chi_{A_k} = \psi_N. \text{ Then } \sup_{\varphi \leq f, \varphi \ \text{simple}} \left\{ \int_a^b \varphi(x)dx \right\} \leq \int_E \chi_{A_k} d\mu \leq \int_E f d\mu \leq M \cdot \mu(E). \)**
\[ \inf_{\psi \geq f, \text{ simple}} \{\int_{a}^{b} \psi(x)dx\} \] and

\[ \int_{E} \psi_N - \varphi_N = \sum \frac{1}{N} \mu(A_k) \leq \frac{1}{N} \mu(E) \to 0 \]

\((\implies)\) We find sequences \(\varphi_N\) and \(\psi_N\) such that they achieve the sup and \text{inf} of the integral, and prove that \(\varphi^* = \sup \{\varphi_N\} = f = \inf \{\psi_N\} = \psi^*\) A.E. \(\blacksquare\)

Then, we can define for this class of functions:

\[ \int_{E} f d\mu \sup_{\varphi \leq f, \text{ simple}} \{\int_{E} \varphi(x)dx\} = \inf_{\psi \geq f, \text{ simple}} \{\int_{E} \psi(x)dx\} \]

Properties:

1. \(\int_{E} cf = c \int_{E} f\)
2. \(\int_{E} (f + g) = \int_{E} f + \int_{E} g\)
3. \(f \geq 0 \implies |\int_{E} f| \leq \int_{E} |f|\)

**Theorem 11** *(Bounded Convergence)* Let \(\{f_n\}_{n=1}^{\infty}\) a sequence of uniformly bounded \(|f_n| \leq M\), measurable functions on \(E\) of finite measure. If \(f_n \to f\) A.E. then:

\[ \int_{E} f d\mu = \lim_{n \to \infty} \int_{E} f_n d\mu \]

**Proof.** Purely by Egorov’s Theorem (Littlewood #3): \(\forall \varepsilon > 0 \exists A \subset E \text{ s.t. } \mu(A) < \varepsilon\) and \(f_n \to f\) unif on \(E \setminus A\) and:

\[ \left| \int_{E} (f - f_n) d\mu \right| \leq \int_{E} |f - f_n| d\mu = \int_{A} |f - f_n| d\mu + \int_{E \setminus A} |f - f_n| d\mu \to 2M \varepsilon \]

\(\blacksquare\)

We observe that if we drop boundedness or finite measure, this not need hold: \(n^2 \chi_{(1/n, 2/n)}\) and \(\chi_{(n, n+1)}\) are examples, respectively.

4. For **Nonnegative Functions**:

For general non-negative functions, we can still use the definition of the integral as the supremum of the integral of simple functions from below if we restrict it to bounded simple functions whose support has finite measure. That is, if we define \(BS(f) = \{\varphi \text{ simple : } \varphi \text{ is simple, bounded and } \mu(supp(f)) < \infty\}\),

\[ \int_{E} f d\mu \sup_{\varphi \in BS(f)} \{\int_{E} \varphi(x)dx\} \]

Properties: For \(f, g\) non-negative:

1. \(\int cf = c \int f\)
2. \( \int (f + g) = \int f + \int g \)

**Proof.** First, since \( \varphi_1 \in BS(f), \varphi_2 \in BS(g) \implies \varphi_1 + \varphi_2 \in BS(f + g) \) then
\[
\int (f + g) \geq \int f + \int g
\]

Given \( \varphi \in BS(f + g) \), we take \( \varphi_1 = \min(\varphi, f) \in BS(f) \) and \( \varphi_2 = \varphi - \varphi_1 \in BS(g) \implies \int (f + g) \leq \int f + \int g \)

3. \( \int f \geq 0, f \geq g \implies \int f \geq \int g \)

**Lemma 12 (Fatou’s Lemma)** Suppose \( \{f_n\} \) a sequence of non-negative measurable functions, \( f_n \to f \) A.E.

\[
\int f \leq \lim_{n \to \infty} \int f_n
\]

**Proof.** It suffices to show \( \int \varphi \leq \lim_{n \to \infty} \int f_n \forall \varphi \in BS(f) \). For any such \( \varphi \), we take \( \varphi_n = \min(\varphi, f_n) \), and so \( \varphi_n \leq \varphi \forall n \). By Bounded Convergence Theorem, \( \varphi_n \to \min(\varphi, f) = \varphi \) and since \( \int \varphi_n \leq f_n \), we get what we wanted.

**Remark 13** Inequality in Fatou’s Lemma can be strict \( (\chi_{(n,n+1)}) \).

**Corollary 14 (Generalization of Fatou’s Lemma)** For an arbitrary \( \{f_n\} \) a sequence of non-negative measurable functions,

\[
\int \lim_{n \to \infty} f_n \leq \lim_{n \to \infty} \int f_n
\]

**Proof.** Apply Fatou to \( g_n = \inf_{k \geq n} \{f_k\} \)

**Corollary 15 (Monotone Convergence Theorem)** Let \( \{f_n\} \) a monotonely increasing sequence of non-negative measurable functions \( f_n \to f \) A.E. Then:

\[
\int f = \lim_{n \to \infty} \int f_n
\]

**Proof.** We apply Fatou’s Lemma to \( f \) and to \( f - f_n \), and we get:

\[
\int f \leq \lim_{n \to \infty} \int f_n \leq \lim_{n \to \infty} \int f_n \leq \int f
\]

5. The Lebesgue Integral (for Measurable functions): We can now define \( f_+ = \max\{f, 0\}, f_- = \max\{-f, 0\} \) non-negative functions. Then \( f = f_+ - f_- \) and \( |f| = f_+ + f_- \). We define:

\[
\int f = \int f_+ - \int f_-
\]
Definition 16 $f \in L^1(\mu)$ (is Lebesgue Integrable) if it is measurable and both $\int f_+ \text{ and } \int f_- < \infty$. Therefore, $f \in L^1(\mu) \iff |f| \in L^1(\mu)$.

Also, $f, g \in L^1(\mu) \Rightarrow f \pm g$, since $f \geq 0$ and $\int f \leq \infty \Rightarrow \mu\{|f = \infty\} = 0$.

Theorem 17 (Lebesgue’s Dominated Convergence Theorem) Let $\{f_n\}$ sequence of measurable functions s.t. $|f_n| \leq g \ \forall n$ for $g \in L^1(\mu)$. If $f_n \to f \ A.E.$ then:

$$\int f = \lim_{n \to \infty} \int f_n$$

Proof. 2 Fatous: For $g - f_n$ and $g + f_n$.

Remark 18 This tells us in the Monotone Convergence Theorem, we can drop monotonicity if we know $f_n \leq \lim_{n \to \infty} f_n \ \forall n$.

Remark 19 Fatou can be generalized to $|f_n| \leq g_n$, $\{g_n\}, g \in L^1(\mu)$ and $\int g_n \to \int g$.

3.3 Convergence in Measure

Definition 20 We say a sequence of measurable functions $\{f_n\}$ converge in measure to $f$, or $f_n \to^m f$ if

$$\mu\{|f_n - f| > \delta\} \to 0 \ \forall \delta > 0$$

1. A.E. convergence on $E$ with $\mu(E) < \infty$ implies convergence in measure (by Egorov / Littlewood #3). A counterexample with $\mu(E) = \infty$ is $f_n = \chi_{(n, \infty)}$.

2. If $\int |f_n - f| \, d\mu \to 0$ then $f_n \to^m f$ (by Markov’s Inequality):

$$\int |f_n - f| \, d\mu \geq \int_{\{|f_n - f| > \varepsilon\}} |f_n - f| \, d\mu \geq \varepsilon \mu\{|f_n - f| > \varepsilon\}$$

3. The limit in measure is unique A.E.: $\{|f - g| > 2\varepsilon\} \subset \{|f - f_n| > \varepsilon\} \cup \{|f_n - g| > \varepsilon\}$

4. The converse is false: If we define $I_{(m,k)} = \left(\frac{k-1}{m}, \frac{k}{m}\right)$, and for $n$ we take $m(n) = \max_{p \in \mathbb{N}} \left\{ \frac{p(p+1)}{2} : \frac{p(p+1)}{2} < n \right\}$, $k(n) = n - m(n)$, $f_n = \chi_{I_{(m(n), k(n))}}(x)$ then this sequence converges in measure, but not pointwise (to 0).

5. We do have the following:

Theorem 21 $f_n \to^m f \iff \exists \{f_{n_k}\}_{k=1}^n \subset \{f_n\}$ subsequence such that $f_{n_k} \to f \ A.E.$
Proof. We take any summable $\varepsilon_k > 0$ s.t. $\varepsilon_k \downarrow 0$. Then, $\exists N_k$ s.t. $n \geq N_k$ 
$\implies \mu(\{|f_n - f| > \varepsilon_k\}) < \varepsilon_k$. Then, we take the subsequence $\{f_{N_k}\}$.

Then, if $E_k = \mu(\{|f_{n_k} - f| > \varepsilon_k\})$. Then $x \in C = \bigcup_{l=1}^{\infty} \bigcap_{k \geq l} (E_k^c)$ $\implies$
$f_{n_k}(x) \to f(x)$

Finally, $\mu(C^c) = \mu(\bigcap_{l=1}^{\infty} \bigcup_{k \geq l} E_k) \leq \sum_{k \geq l} \mu(E_k) < \sum_{k \geq l} \varepsilon_k \to 0$.

5. If $\mu(E) < \infty$ and $f_n \to^* f$ $\iff$ every subsequence of $\{f_n\}$ in turn has a
subsequence that converges $A.E.$

6. Substituting convergence $A.E.$ with $f_n \to^* f$, Fatou’s Lemma, Monotone
Convergence and LDCT remain valid.

4 Differentiation and Integration

The main aim of this section is to generalize the two fundamental theorems of
calculus to a larger class of functions using the Lebesgue Integral:

\[
\begin{align*}
\text{FTC1} & : \quad \int_a^b f'(t) dt = f(b) - f(a) \\
\text{FTC2} & : \quad \frac{d}{dx} \int_a^x f(t) dt = f(x)
\end{align*}
\]

The answers we will find are the following: FTC1 holds if $f$ is $AC$ and FTC2 holds $A.E.$ if $f \in L^1(\mu)$.

4.1 Second FTC: Wiener Covers, H-L Maximal Function,
Lebesgue Set

Theorem 22 Let $f \in L^1([a, b])$, then for $A.E.$ $x$ in $[a, b]$,

\[
\frac{d}{dx} \int_a^x f(t) dt = f(x)
\]

Here Prof. Gunturk took the following tortuous yet illustrative path:

Lemma 23 (Wiener Covering Lemma): Let $\{I_\lambda\}_{\lambda \in \Lambda}$ collection of open
intervals in $[a, b]$. Let $U = \bigcup_{\lambda \in \Lambda} I_\lambda$. Then $\exists\{\lambda_1, \ldots, \lambda_n\}$ s.t. $\{I_{\lambda_i}\}_{i=1}^n$ is pairwise
disjoint and:

\[
\sum_{i=1}^{n} |I_{\lambda_i}| \geq \frac{\mu(U)}{4}
\]
**Definition 24** Let \( f \in L^1([a,b]) \). We define \( A_h f(x) = \frac{1}{2h} \int_{x-h}^{x+h} f(t) \, dt \), which is a continuous function. Then, the **Hardy-Littlewood Maximal Function** is defined as:

\[
M f(x) = \sup_{h > 0} \{ A_h f(x) \}
\]

**Lemma 25** (Maximal Inequality) Let \( f \in L^1([a,b]) \). Then, \( \forall \lambda > 0 \),

\[
\mu(\{ M f > \lambda \}) \leq \frac{C \int |f|}{\lambda}
\]

**Proof.** \( E_\lambda = \{ M f > \lambda \} = \bigcup_{h > 0} \{ A_h |f| (x) > \lambda \} \) union of open sets. So, for each \( x \in E_\lambda \) \( \exists h_x > 0 \) s.t. \( A_{h_x} |f| (x) > \lambda \). I take the covering \( \{ B_{h_x}(x) \}_{x \in E_\lambda} \) and by Wiener’s Covering Lemma, I can find a subcover such that:

\[
\mu(E_\lambda) \leq 8 \sum_{i=1}^{N} h_{x_i} \leq \sum_{i=1}^{N} \frac{4}{\lambda} \int_{B_{h_i}(x)} |f| \leq \frac{4}{\lambda} \int |f|
\]

\[\blacksquare\]

So now, the strategy is to show \( A_h f \rightarrow f \) as \( h \rightarrow 0 \) by approximating \( f \) by a continuous function and then controlling the error with the maximal function:

**Proposition 26** Let \( f \in L^1(\mathbb{R}) \). Then, \( \forall \varepsilon > 0 \ \exists g \in C(\mathbb{R}) \) with compact support such that \( \int |f - g| < \varepsilon \)

**Proof.** This can either be shown with results of test functions and convolution, or by a 3\( \varepsilon \) proof: 1) truncating the domain, 2) making the function bounded and 3) approximating with a continuous function. \[\blacksquare\]

**Proposition 27** Let \( f \in L^1(\mathbb{R}) \). Then, \( A_h f(x) \rightarrow f(x) \) as \( h \rightarrow 0 \) A.E.

**Proof.**

\[
\lim_{h \to 0} |A_h f - f| \leq \lim_{h \to 0} |A_h (f - g)| + \lim_{h \to 0} |A_h g - g| + |(f - g)|
\]

\[
\Rightarrow \lim_{h \to 0} |A_h f - f| \leq M |f - g| + |(f - g)|
\]

Using the maximal inequality, this means that \( E^\alpha = \{ \lim_{h \to 0} |A_h f - f| > \alpha \} \subseteq \{|f - g| > \alpha/2\} \cup \{M |f - g| > \alpha/2\}, \mu(E^\alpha) \leq C\varepsilon/\alpha \forall \varepsilon. \[\blacksquare\]

**Definition 28** (**Lebesgue Set**) Let \( f \in L^1(\mu) \). The **Lebesgue Set of** \( f \) is the set of points such that \( \frac{1}{2h} \int_{x-h}^{x+h} |f(t) - f(x)| \, dt \rightarrow 0 \) as \( h \rightarrow 0 \).

**Theorem 29** Let \( f \in L^1(\mu) \), then almost every \( x \) is in the **Lebesgue Set**.

**Proof.** For any \( r \in \mathbb{R} \), by last proposition,

\[
\lim_{h \to 0} \frac{1}{2h} \int_{x-h}^{x+h} |f(t) - r| \, dt = |f(x) - r| \ \forall x \in E^\alpha_r \text{ with } \mu(E^\alpha_r) = 0
\]
Then, $E = \bigcup_{r \in \mathbb{Q}} E_r$, $\mu(E) = 0$. For $x \in E^c$, since $\mathbb{Q}$ is dense in $E^c$, we obtain our result. Finally,

$$\frac{1}{h} \int_x^{x+h} |f(t) - f(x)| \, dt \leq \frac{1}{h} \int_{x-h}^{x} |f(t) - f(x)| \, dt \to 0$$

\[ \blacksquare \]

### 4.2 First FTC: AC, BV and Vitali Cover

Now, we want to know when the first fundamental theorem holds. We could admit weak (distribution derivatives), and then this statement would be true for a really large space of functions ($H^1(\mathbb{R})$). If we insist on the derivative being a function, then this works if $f$ is Absolutely Continuous:

**Definition 30** $f : [a, b] \to \mathbb{R}$ is absolutely continuous if $\forall \varepsilon > 0 \exists \delta > 0$ s.t. for all non-overlapping intervals $\{(a_i, b_i)\}_{i=1}^k$ such that $\sum_{i=1}^k |b_i - a_i| < \delta$ then

$$\sum_{i=1}^k |f(b_i) - f(a_i)| < \varepsilon$$

**Proposition 31** Let $f \in L^1(\mu)$ then $F(x) = \int_a^x f(t) \, dt$ then $F$ is absolutely continuous.

**Proof.** We prove that, for $f \in L^1(\mu)$, $\forall \varepsilon > 0 \exists \delta > 0$ s.t. $\forall A$ s.t. $\mu(A) < \delta \implies \int_A |f| < \varepsilon$ (using the sequence $f_n(x) = f(x)\chi_{[f \leq n]}(x)$ ) \[ \blacksquare \]

**Definition 32** Let $f : [a, b] \to \mathbb{R}$, $P = \{a = t_0 < t_1 < \ldots < t_n = b\}$ a partition of $[a, b]$. Then, if we define $V(P, f) = \sum_{i=1}^n |f(t_i) - f(t_{i-1})|$, $V_+(P, f) = \sum_{i=1}^n [f(t_i) - f(t_{i-1})]_+$ and $V_-(P, f) = \sum_{i=1}^n [f(t_i) - f(t_{i-1})]_-$ then:

$$TV^f_a(f) = \sup_P \{V(P, f)\}$$

$$PV^f_a(f) = \sup_P \{V_+(P, f)\}$$

$$NV^f_a(f) = \sup_P \{V_-(P, f)\}$$

Then, by definition, $TV^f_a(f) = PV^f_a(f) + NV^f_a(f)$, $TV^f_a(f) \geq 0$ if $f$ and $TV^f_a(f) = TV^f_a(f) + TV^f_c(f) \forall c \in [a, b]$.

**Definition 33** A function $f : [a, b] \to \mathbb{R}$ is of Bounded Variation ($\in BV([a, b])$) $\iff$ $TV^f_a(f) < \infty$. Also, a norm in $BV([a, b])$ is given by $|f|_{BV([a, b])} = TV^f_a(f) + |f(a)|$. 

12
Lemma 34 A function \( f \in BV([a,b]) \), then:

\[
f(b) - f(a) = PV^b_a(f) - NV^b_a(f)
\]

Theorem 35 A function \( f \in BV([a,b]) \) \iff \( f = g - h \) where \( g \) and \( h \) are monotone increasing

Proof. \((\Rightarrow)\) \( f \in BV([a,b]) \) then, by the previous lemma, \( f(x) - f(a) = PV^x_a(f) - NV^x_a(f) \). Then \( g(x) = PV^x_a(f) \) and \( h(x) = NV^x_a(f) - f(a) \) do the trick.

\((\Leftarrow)\) \( f = g - h \) \implies \( V(P, f) \leq V(P, g) + V(P, h) = g(b) + h(b) - g(a) - h(a) \) \( \forall P \) partition of \([a,b]\). Then, \( TV^b_a(f) \leq TV^b_a(g) + TV^b_a(h) = g(b) + h(b) - g(a) - h(a) < \infty \) (which, by the way, could have been established by triangle inequality for \( \|f\|_{BV([a,b])} \)).

Theorem 36 \( AC([a,b]) \subset BV([a,b]) \)

Proof. Given \( f \in AC([a,b]) \), we take a partition \( P \) such that \( \|P\| < \delta(1) \). Then, we block the resulting intervals in blocks s.t. each spans an interval of length between \( \delta/2 \) and \( \delta \). Then, the number of blocks is less than \( \left\lceil \frac{2(b-a)}{\delta} \right\rceil \), and on each block, \( \sum |f(t_i) - f(t_{i-1})| \leq \varepsilon = 1 \). This implies \( V(P, f) \leq \varepsilon \left\lceil \frac{2(b-a)}{\delta} \right\rceil = \left\lceil \frac{2(b-a)}{\delta} \right\rceil \) \( \forall \|P\| < \delta \). This means \( TV^b_a(f) \leq \left\lceil \frac{2(b-a)}{\delta} \right\rceil < \infty \). \( \blacksquare \)

Finally, the fundamental theorem follows from two important facts: first, that if \( f \) is monotonic on \([a,b]\), then it is almost everywhere differentiable and its derivative is integrable. Since any bounded variation function is the difference of two monotone increasing functions, this fact is also true for \( BV([a,b]) \). Since \( AC \subset BV \), we can use this and the second fundamental theorem to prove our result.

Definition 37 Let \( E \) be any set, and \( \Upsilon \) a collection of nondegenerate intervals. We say \( \Upsilon \) is a Vitali Covering of \( E \) if:

\[
\forall \varepsilon > 0, \forall x \in E \ \exists I \in \Upsilon \ s.t. \ x \in I \text{ and } |I| < \varepsilon
\]

This covering allows us to refine arbitrarily around any point in \( E \).

Lemma 38 (Vitali) Let \( E \subset [a,b] \) and \( \Upsilon \) a Vitali cover of \( E \). Then, \( \forall \varepsilon > 0, \exists \) a finite disjoint subcollection \( \{I_k\}_{k=1}^N \subset \Upsilon \) s.t.

\[
\mu^*(E \setminus \bigcup_{i=1}^N I_i) < \varepsilon
\]
In particular, there
Then:

\[ f \] is differentiable everywhere, we use a more classical approach using "derivates":

Proof. I work with closed intervals, and follow a greedy approach. Having selected the first \( n \) intervals, \( \lambda_n = \text{sup}\{ |I| : I \cap I_i = \emptyset, I \in \mathcal{I} \} \), and we choose \( I_{n+1} \) such that its length is greater than \( \lambda_n/2 \).

Unless the process terminates at a finite \( N \), we have an infinite sequence of disjoint intervals. Since they all lie in \([a, b] \), \( \forall \varepsilon > 0 \ \exists N(\varepsilon) \) s.t. \( \sum_{i=N+1}^{\infty} |I_i| < \frac{\varepsilon}{5} \)
then:

\[ E \setminus \bigcup_{i=1}^{N} I_i \subset \bigcup_{i=N+1}^{\infty} (5I_i) \]

\[ \blacksquare \]

**Proposition 39** \( f \in AC([a, b]) \) and \( f' \equiv 0 \ A.E. \implies f \) is constant.

Proof. \( \forall \varepsilon > 0 \) let \( \delta(\varepsilon) \) of the definition of absolute continuity. It suffices to show \( f(a) = f(b) \ \forall a, b \). Then \( f'(x) = 0 \implies \)

\[ \frac{|f(x+h) - f(x)|}{h} < \eta \ \forall \text{ sufficiently small } \eta \]

In particular, there \( \exists \) arbitrarily small values of \( h \) s.t. \( |f(x+h) - f(x)| < \eta h \). Then all such \([x, x+h]\) form a Vitali cover of the set \( E = \{ x : f'(x) = 0 \} \).

By Vitali’s Lemma, for \( \delta(\varepsilon) \) we choose the finite disjoint subcollection \([x_i, y_i]\),
\( x_1 < y_1 < x_2 < \ldots < y_N, y_0 = a \) and \( x_{N+1} = b \). Then \( \sum |x_k - y_{k-1}| < \delta \) and:

\[ \sum |f(y_i) - f(x_i)| < \eta \sum |y_i - x_i| \leq \eta(b-a) \]

\[ \sum |f(x_k) - f(y_{k-1})| < \varepsilon \]

And so, \( |f(b) - f(a)| < V(P, f) = \varepsilon + \eta(b-a) \). \[ \blacksquare \]

Although there are more abstract ways to prove a montonous function is differentiable everywhere, we use a more classical approach using "derivates":

**Definition 40** We define the four one-sided derivates, which exist \( \forall x \), as follows:

\[ D^+ f(x) = \lim_{h \to 0^+} \frac{f(x+h) - f(x)}{h} \], \( D^+_f(x) = \lim_{h \to 0^+} \frac{f(x+h) - f(x)}{h} \)

\[ D^- f(x) = \lim_{h \to 0^+} \frac{f(x) - f(x-h)}{h} \], \( D^- f(x) = \lim_{h \to 0^+} \frac{f(x) - f(x-h)}{h} \)

Clearly, \( D^+ f(x) \geq D^+_f(x) \) and \( D^- f(x) \geq D^- f(x) \ \forall x \), and \( f' \ast(x) \ \exists \iff D^+ f(x) = D^+_f(x) = D^- f(x) = D^- f(x) \). If the one sided derivates coincide, we call them the right or left hand side derivatives.

**Proposition 41** If \( f \) is monotone increasing on \([a, b]\), then it is differentiable A.E. and its derivative is integrable.
It then suffices to show \( \mu(E_{u,v}) = 0 \) \( \forall u, v \in \mathbb{Q} \). For \( x \in E_{u,v} \), we have that there exists an arbitrarily small \( h > 0 \) s.t. \( f(x) - f(x - h) < vh \) (intervals \( (x - h, x] \subset \mathcal{T}_1 \)) and \( f(x + h) - f(x) > uh \) (intervals \( [x, x + h) \subset \mathcal{T}_2 \)). For a given \( E_{u,v} \) and \( \varepsilon > 0 \), \( \exists \Omega \) open such that \( E_{u,v} \subset \Omega \) and \( \mu(\Omega) \leq \mu^*(E_{u,v}) + \varepsilon \). We have that \( \mathcal{T}_1 \) is a Vitali cover of \( E_{u,v} \), and hence, \( \exists \{[x_i - h_i, x_i]\} \) such that:

\[
\mu^*(E_{u,v} \setminus \bigcup_{i=1}^N [x_i - h_i, x_i]) < \varepsilon
\]

\[
\mu^*(E_{u,v}) < \mu^*(E_{u,v} \cap \bigcup_{i=1}^N [x_i - h_i, x_i]) + \varepsilon
\]

\[
\sum_{i=1}^N f(x_i) - f(x_i - h_i) < v \sum_{i=1}^N h_i < v(\mu^*(E_{u,v}) + \varepsilon)
\]

Let \( A_{u,v} = E_{u,v} \cap \bigcup_{i=1}^N [x_i - h_i, x_i] \). Now, we consider the intervals in \( \mathcal{T}_2 \) as a Vitali cover of \( A_{u,v} \). Again, we obtain \( \{[y_j, y_j + h_j]\} \) s.t. each interval is inside one of the intervals \([x_i - h_i, x_i]\) and:

\[
\mu^*(A_{u,v} \cap \bigcup_{j=1}^M [y_j, y_j + h_j]) > \mu^*(A_{u,v}) - \varepsilon > \mu^*(E_{u,v}) - 2\varepsilon
\]

\[
v(\mu^*(E_{u,v}) + \varepsilon) > \sum \left| f(y_j) - f(y_j + h_j) \right| > u(\mu^*(E_{u,v}) - 2\varepsilon)
\]

\[
\implies (v - u)\mu^*(E_{u,v}) = 0 \implies \mu^*(E_{u,v}) = 0
\]

Hence, the derivatives are equal \( A.E. \). Now, \( g_n(x) = \frac{f(x + h_n) - f(x)}{h_n} \) with \( h_n \downarrow 0 \) converges \( A.E. \) to \( f'(x) \) and since \( f \) is monotonic, \( g_n \geq 0 \). By Fatou’s Lemma,

\[
\int_{[a,b]} f'(x) \, d\mu \leq \lim_{n \to \infty} \int_{[a,b]} g_n
\]

First, we show this limit is finite

\[
\int_{[a,b]} g_n = \frac{1}{h_n} \left( \int_a^{b+h_n} f - \int_a^{a+h_n} f \right) \leq f(b) - f(a) < \infty
\]

This proves the derivative is integrable: however, this is an equality only if \( f \in AC([a, b]) \).
Finally, with the help of these two propositions, we prove the result we have been looking for:

**Theorem 42** Let $f \in AC([a, b])$. Then $\int_a^b f'(t)dt = f(x) - f(a)$ $\forall x \in [a, b]$.

**Proof.** We first define $\varphi(x) = f(a) + \int_a^x f'(t)dt$, and our goal is to prove $\varphi \equiv f$. Since $f' \in L^1([a, b])$ by the last proposition and Lemma ( ), $\varphi$ is differentiable and by the second FTC, $\varphi' \equiv f' \ A.E.$ Now, $\varphi - f \in AC([a, b])$ and $(\varphi - f)' \equiv 0 \ A.E.$, so by the first proposition we have that $\varphi - f \equiv C$ constant. Evaluating on $a$, $C = \varphi(a) - f(a) = 0$. □

**Remark 43** This theorem tells us that a function is absolutely continuous $\iff$ it is the indefinite integral of its derivative.

### 4.3 Convex Functions and Jensen’s Lemma

**Definition 44** $\phi: (a, b) \to \mathbb{R}$ is convex if $\forall x < y \in (a, b)$, $\forall \lambda \in [0, 1]$,

$$\phi(\lambda x + (1 - \lambda)y) \leq \lambda\phi(x) + (1 - \lambda)\phi(y)$$

That is, the graph of $\phi$ always lies below the segment joining $(x, \phi(x))$ and $(y, \phi(y))$. If $-\phi$ is convex, we will say $\phi$ is concave.

**Definition 45** An equivalent definition for convexity is the following: $\phi$ is convex if $\forall x < t < y$,

$$\frac{\phi(t) - \phi(x)}{t - x} \leq \frac{\phi(y) - \phi(x)}{y - x} \leq \frac{\phi(y) - \phi(t)}{y - t}$$

The main results that we are interested in are:

1. A convex function is Lipschitz continuous on every closed subinterval of $(a, b)$
2. A convex function is differentiable everywhere except possibly on countably many points
3. $\phi''$ exists and is $> 0$ $A.E.$ (because $\phi'$ is monotonic).
4. **Jensen’s Inequality**: $\phi$ a convex function on $\mathbb{R}$ and $f \in L^1([0, 1])$ then:

$$\phi\left(\int_0^1 f(t)dt\right) \leq \int_0^1 \phi(f(t))dt$$

**Definition 46** $\phi$ is Lipschitz continuous on $[a, b]$ if $\exists L \in [0, \infty)$ s.t. $|\phi(x) - \phi(y)| \leq L |x - y|$ $\forall x, y \in [a, b]$. This condition is stronger than absolute continuity, since it also implies the derivative is bounded.
Proof. (1) $\phi$ convex in $(a, b), [c, d] \subset (a, b)$. Then, if we pick $e \in (a, c)$ and $f \in (d, b)$. By the alternative definition of convexity, we know:

$$\frac{\phi(c) - \phi(e)}{c - e} \leq \frac{\phi(y) - \phi(x)}{y - x} \leq \frac{\phi(f) - \phi(d)}{f - d} \quad \forall x, y \in [c, d]$$

Now, if we let $L = \max\left\{ \frac{\phi(c) - \phi(e)}{c - e}, \left| \frac{\phi(f) - \phi(d)}{f - d} \right| \right\}$, then we have our Lipschitz condition.

(2) We know that right hand and left hand derivatives exist everywhere for a convex function. (?)

(3) It follows from (2), since the derivative exists $A.E.$ and is monotonic. ■

Definition 47 $\phi(x_0) + m(x - x_0)$ is a supporting line for $\phi$ at $x_0$ if $\phi(x_0) + m(x - x_0) \leq \phi(x) \quad \forall x$.

Proof. (Jensen’s Inequality) (2) implies that every convex function $\phi$ has a supporting line at every point (and $\phi'_-(x_0) \leq m_{x_0} \leq \phi'_+(x_0) \forall x_0$).

Now, given a supporting line, we set $x = f(t)$ and integrate:

$$\phi(x_0) + m_{x_0} \int_0^1 f(t) - m_{x_0} x_0 \leq \int_0^1 \phi(f(t))dt$$

Finally, we choose $x_0 = \int_0^1 f(t)dt$, and:

$$\phi(\int_0^1 f(t)dt) + m_{x_0} \int_0^1 f(t) - m_{x_0} \int_0^1 f(t)dt = \phi(\int_0^1 f(t)dt) \leq \int_0^1 \phi(f(t))dt$$

■

5 $L^p$ spaces

Definition 48 We define, for $p \in (0, \infty) \ L^p(\mathbb{R}) = \{ f \text{ measurable} : |f|^p \in L^1(\mathbb{R}) \}$. For $p \in [1, \infty)$, the set of equivalence classes $A.E.$ $(L^p(\mathbb{R}), \| \|_p)$ is a normed vector space with the norm:

$$\|f\|_p = \left( \int |f(x)|^p \, d\mu \right)^{\frac{1}{p}}$$

Definition 49 We define $\text{ess. sup}(f) = \inf \{ M : f(x) \leq M \ A.E. \}$ the smallest bound $A.E.$ and $L^\infty(\mathbb{R}) = \{ f \text{ measurable} : \text{ess. sup}(f) < \infty \}$. This is a normed vector space with the norm $\|f\|_\infty = \text{ess. sup}(f)$.

Proposition 50 $\|f\|_\infty \leq L \iff |f| \leq L \ A.E.$
Theorem 51 (Minkowski Inequality) If $p \in [1, \infty]$, $\|f + g\|_p \leq \|f\|_p + \|g\|_p$ (this inequality is reversed for $p \in (0, 1)$). If $p > 1$, equality is obtained $\iff \exists C > 0$ s.t. $f = Cg$. For $p = 1$ (since the space is not strictly convex) equality is attained if $f$ and $g$ have the same sign.

Proof. If $f$ or $g$ of norm zero, then the inequality is trivial. Otherwise, $	ilde{f} = \frac{f}{\|f\|_p + \|g\|_p}$, $\tilde{g} = \frac{g}{\|f\|_p + \|g\|_p}$ and $\lambda = \frac{\|f\|_p}{\|f\|_p + \|g\|_p}$. By convexity of $t^p$,

$$\left| \tilde{f} + \tilde{g}\right|^p \leq \left( \left| \tilde{f} \right| + \left| \tilde{g} \right| \right)^p$$

$$= \left( \frac{|f|}{\|f\|_p} - \lambda + \frac{|g|}{\|g\|_p} (1 - \lambda) \right)^p$$

$$\leq \left( \frac{|f|^p}{\|f\|_p^p} \lambda + \frac{|g|^p}{\|g\|_p} (1 - \lambda) \right)^p$$

Integrating on both sides, $\left| \tilde{f} + \tilde{g}\right| \leq 1$ and the result follows. Conditions for equality follow from analyzing the two steps in which there is an inequality. ■

Proposition 52 (Young’s Inequality) $a, b \geq 0$, $p \geq 1$ and $\frac{1}{p} + \frac{1}{q} = 1$ then:

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

If $p \in (0, 1)$, the inequality is reversed. Equality in both cases is achieved $\iff a^p = b^q$.

Theorem 53 (Hölder Inequality) For $f \in L^p$, $g \in L^q$ for $\frac{1}{p} + \frac{1}{q} = 1$ then $fg \in L^1$ and:

$$\int |fg| \leq \|f\|_p \|g\|_q$$

If $p \in (0, 1)$, the inequality is reversed. Equality in both cases is achieved $\iff |f|^p \|g\|_q = |g|^q \|f\|_p^p$.

Proof. If $q < \infty$, We use Young’s Inequality with $a = \frac{|f(x)|}{\|f\|_p}$, $b = \frac{|g(x)|}{\|g\|_q}$ and integrate on both sides. The case $q = \infty$ is trivial, since $|g(x)| \leq \|g\|_\infty$. ■

Corollary 54 If $\mu(\Omega) < \infty$, then for $p_1 \leq p_2$, $L^{p_2}(\Omega) \subset L^{p_1}(\Omega)$ ($L^\infty(\Omega) \subset \subset L^1(\Omega)$). If $\Omega$ is of infinite measure, this fails completely.

Definition 55 We define, for $p \in (0, \infty)$ the space $l^p(\mathbb{N}) = \{x_n\}_1^\infty \in \mathbb{R}^\omega$ : $\sum_{n=1}^\infty |x_n|^p < \infty$. The space $(l^p, \|\cdot\|_p)$ is a normed vector space equipped with the norm:

$$\|\{x_n\}_1^\infty\|_p = \left( \sum_{n=1}^\infty |x_n|^p \right)^{\frac{1}{p}}$$
This space can be isometrically embedded onto a closed subspace of $L^p(\mathbb{R})$ using the linear mapping: $\phi(\{x_n\}) = f(\{x_n\}) = \sum_{n=1}^{\infty} x_n x_{n-1, n}(x)$.

**Definition 56** In an analogous fashion, we define $l^\infty(\mathbb{N}) = \{\{x_n\}_1^\infty \in \mathbb{R}^\infty : \sup_n |x_n| < \infty\}$. This space is a normed vector space equipped with the norm $\|\{x_n\}\|_\infty = \sup_n |x_n|$. It can also be isometrically embedded onto a closed subspace of $L^\infty(\mathbb{R})$ using the same mapping $\phi$.

**Remark 57** Using the mapping $\phi$, we can easily check that Minkowski and Hölder Inequalities also hold for the $p$ spaces. This is also true because these are $L^p$ spaces on $\mathbb{N}$ corresponding to the counting measure. We will see that most of the results we prove for the $L^p$ spaces will immediately be applicable to the $L^p$s.

A very important property of a topological vector space is completeness: that all Cauchy sequences are convergent. If a normed vector space is complete, we will say it is a Banach space. To prove that $L^p$ spaces are all Banach spaces, we first state a characterization of Banach spaces in terms of summability of series.

**Definition 58** Given $\{x_n\}_{n=1}^\infty \in X$ normed vs, we say the corresponding series is summable if $s_n = \sum_{k=1}^{n} x_k$, $s_n \to s$. We say it is absolutely summable if $\sum_{n=1}^{\infty} ||x_n|| < \infty$.

**Theorem 59** A normed vector space $(X, ||\cdot||)$ is complete $\iff$ every absolutely summable series is summable.

**Proof.** ($\Rightarrow$) Let $\{x_n\}$ be absolutely summable. Then $s_m - s_k \leq \sum_{m+1}^{k} ||x_n|| \to m-\infty 0$. Thus, $s_m$ is a Cauchy sequence, and by our hypothesis, it is convergent $\implies x_n$ is summable.

($\Leftarrow$) Let $\{x_n\}$ be a Cauchy sequence in $X$. Then, $\exists m_k \text{ s.t. } ||x_{m+1} - x_n|| < 2^{-k}$. We call $y_k = x_{m+1} - x_n$ and $y_k$ is absolutely summable. Now, by our hypothesis, $y_k$ is summable, and $\sum_{k=1}^{\infty} y_k = \lim_{k \to \infty} \{x_{m+1} - x_n\} = s$. Thus, $\{x_n\}$ has a convergent subsequence, and hence, the whole sequence converges.

**Theorem 60** (Riesz-Fischer) For $p \in [1, \infty]$, $L^p$ spaces are Banach spaces.

**Proof.** For both cases $p = \infty$ and $p < \infty$ we show an absolutely summable series in $L^p$ is summable:

($p = \infty$) Let $\{f_n\}_1^\infty \subset L^p$, $s_n = \sum_{k=1}^{n} f_k$. Then, we know $|s_n - s_m| (x) \leq \sum_{k=m+1}^{n} |f_k(x)|$. Now, for each $k$, $\exists E_k \text{ s.t. } \mu(E_k) = 0$ and $|f_k(x)| \leq ||f_k||_\infty$. $\forall x \in E_k$. Then, on $E = \cap_{k=1}^{\infty} E_k$:

$$|s_n - s_m| (x) \leq \sum_{k=m+1}^{n} ||f_k||_\infty \to m-\infty 0$$
So, $s_n(x)$ is a Cauchy sequence $\forall x \in E \implies \exists s$ s.t. \( s(x) = \lim_{n \to \infty} s_n(x) \) on $E$. Finally,

$$\|s_n - s\|_\infty \leq \sum_{k=m+1}^{\infty} \|f_k\|_\infty \rightarrow_{m \to \infty} 0$$

And so, the series is summable in $L^\infty$.

($p < \infty$) Now we also define $a_m(x) = \sum_{n=1}^{m} |f_n|(x)$ series of absolute values. Now, we know $\|a_m\|_p \leq \sum_{k=1}^{m} \|f_k\|_p \leq \sum_{k=1}^{\infty} \|f_k\|_p < \infty$. Also, it is monotonically increasing so, by monotone convergence theorem $\exists a(x)$ limit and $\int a(x)^p dx = \lim_{m \to \infty} \int a_m(x)^p dx \leq C^p < \infty$. Therefore, $a \in L^p$ and by the same argument, $s_m(x)$ has a pointwise limit $s(x)$ and by Lebesgue Dominated Convergence Theorem ($s_m(x)$ is dominated by $a(x)$),

$$\int |s_m(x) - s(x)|^p \rightarrow \int \lim_{n \to \infty} |s_n(x) - s(x)|^p dx = 0$$

Finally, when studying any normed vector space, we can often identify the space of linear, continuous functionals (its topological dual) with some space we can work with. This is sometimes achieved through some kind of representation theorem, and then allows us to deduce properties of the space through the study of its dual.

**Theorem 6.1 (Riesz Representation Theorem for $L^p$ spaces)** For $p \in [1, \infty)$, any $l \in (L^p)^*$ has a unique representer $g_l \in L^q$ s.t. $l(f) = \int f g_l$ $\forall f \in L^p$. The mapping $R : (L^p)^* \rightarrow L^q$ such that $R(l) = g_l$ is a linear isometry: in particular, $\|l\|_* = \|g_l\|_q$.

1. **We prove it for $L^p([a, b])$:**
   (i) All $g$ in $L^q$ work: First, if $g \in L^q$ then $l_g$ given by $l_g(f) = \int f g$ is a continuous linear functional, since, by Holder: $|l_g(f)| \leq \|g\|_q \|f\|_p$. This also tells us that $\|l_g\|_* \leq \|g\|_q$. However, evaluating this functional on $f = |g|^{q-1} \text{sgn}(g) \in L^p$, we find that in fact $\|l_g\|_* = \|g\|_q$.
   (ii) Characteristic Functions: Now, let $l \in (L^p)^*$. For $l_g$ as before, $l_g(\chi_{[a, t]}) = \int_0 t g(t) dt = G(t)$, and we can identify $g$ with $G'$. Now, given our arbitrary $l$, we again consider $G_l(t) = l(\chi_{[a, t]})$ and we show $G_l$ is $AC$, and $g_l = G'_l$.

   Now, given $\{[a_i, b_i]\}_{i=1}^N$ non-overlapping intervals, we consider $f = \sum_{i=1}^{N} c_i \chi_{[a_i, b_i]} = \sum_{i=1}^{N} c_i (\chi_{[a, b_i]} - \chi_{[a, a_i]})$ $A.E.$ Then:

$$l(f) = \sum_{i=1}^{n} c_i [G(b_i) - G(a_i)]$$
Picking \( c_i = sgn(G(b_i) - G(a_i)) \), we now have:

\[
 l(f) = \sum_{i=1}^{n} |G(b_i) - G(a_i)| \leq \|l\|_* \|f\|_p = \|l\|_* \left( \sum_{i=1}^{N} |b_i - a_i| \right)^{\frac{1}{p}}
\]

So, for a given \( \varepsilon \) we can take \( \delta = (\varepsilon / \|l\|_*)^p \). \( G_t \) is \( AC \) and \( g_t = G'_t \in L^1 \) by TFC1.

Also, we now know \( l(\chi_{[a,t]}) = G(t) = \int_{a}^{t} g_t = \int \chi_{[a,t]} g_t \). By linearity of \( l \) and the integral, immediately this implies \( l(\psi) = \int \psi g_t \forall \psi \) step functions.

(iii) From step to \( L^\infty \): Now, our claim now is that \( l(f) = \int f g_t \forall f \in L^\infty([a,b]) \). We have that:

\[
|l(f) - \int f g| = |l(f - \psi) - \int (f - \psi) g| \leq \|l\|_* \|f - \psi\|_p + \int |f - \psi| |g|
\]

Now, for \( \varepsilon > 0 \exists \delta(\varepsilon) \) s.t. \( \forall E \) with \( \mu(E) < \delta, \int_E |g| < \varepsilon \). Also, for this \( \delta, \exists \psi_\delta \) s.t. \( |\psi - f| < \delta \) except on a set \( A \) with \( \mu(A) < \delta \) and \( |\psi| \leq \|f\|_\infty \). Then,

\[
\|f - \psi\|_p \leq \int_A |f - \psi|^p + \int_{A^c} |f - \psi|^p \leq \delta^p (b-a) + 2 \|f\|_\infty \delta \leq C_1(f,a,b) \varepsilon
\]

\[
\int |f - \psi| |g| \leq \int_{A^c} \delta |g| + \int_{A} 2 \|f\|_\infty |g| \leq \delta \|g\|_1 + 2 \|f\|_\infty \varepsilon \leq C_2(g,f,a,b) \varepsilon
\]

So, \( l(f) = \int f g \forall f \in L^\infty([a,b]) \).

(iv) We prove \( g_t \in L^q \) and \( \|g_t\|_q = \|l\|_* : \)

Case: \( q < \infty \) Let \( g_n = g \chi_{[g \leq n]} \) and \( f_n = \sgn(g) \chi_{[g \leq n]} \). Then \( \int f_n g_n = \int |g|^{q-1} \sgn(g) \chi_{[g \leq n]} |g| = \|g\|_q \). Hence, \( \|g_n\|_q \leq \|l\|_* \).

Case: \( q = \infty \) For \( \varepsilon > 0 \), \( |g| \leq \|g\|_\infty - \varepsilon \) on \( A \), for some \( \mu(A) > 0 \). Then \( f = \sgn(g) \chi_A \). Then \( f \) is bounded, \( \|f\|_1 = 1 \) and \( \int f g = \frac{\sgn(g)}{\mu(A)} \int_A |g| \geq \|g\|_\infty - \varepsilon \).

So, \( \|g\|_\infty \leq \|l\|_* \).

(v) We show this for all \( f \in L^p : |l(f) - \int f g| < \|l\|_* \|f - f_n\|_p + \int |f - f_n| |g| \) with \( f_n \) bounded and \( f_n \to f \) in mean \( p \).

(vi) Extension to \( \mathbb{R} \): Let \( l \in (L^p(\mathbb{R}))^* \). For each \( N \), we find \( g_N \) on \([-N,N]\) such that \( l(f) = \int f g_N \forall f \in L^p([-N,N]) \), and also, by uniqueness of the \( g' \)'s, \( g_{N+1} \equiv g_N \forall N \).

So, we have a consistent sequence of \( g_N \) which converge to \( g \) measurable, and \( \|g_N\|_q \leq \|l\|_* \). By monotone convergence theorem, \( g \in L^q(\mathbb{R}) \) and by continuity, \( l(f) = \int f g \forall f \in L^p(\mathbb{R}) \).