MA5205 GRADUATE ANALYSIS I

1. Background and Review

Notation
$\mathbb{N}$: natural numbers, $1, 2, 3, \ldots$
$\mathbb{Z}$: integers
$\mathbb{Q}$: rational numbers
$\mathbb{R}$: real numbers

1.1. **Bounds. Supremum Axiom.** $\limsup$ **and** $\liminf$**.**

1. If $\emptyset \neq A \subseteq \mathbb{R}$, a number $x \in \mathbb{R}$ is an **upper bound (lower bound)** of A if $a \leq x$ ($a \geq x$) for all $a \in A$.
A set A is **bounded above (below)** if it has an upper bound (lower bound).
A set that is both bounded above and bounded below is **bounded**.
A function $f : \Omega \to \mathbb{R}$ is **bounded (above, below)** if $f(\Omega) = \{f(\omega) : \omega \in \Omega\}$ is bounded (above, below).

2. A number x is the **supremum (= least upper bound)** [**infimum (= greatest lower bound)**] of A if (i) x is an upper bound [lower bound] of A and (ii) $x \leq u$ for any upper bound u of A [$x \geq l$ for any lower bound l of A].
Write $\sup A$ and $\inf A$ for the supremum and infimum of A respectively.

3. **Supremum Axiom (= Least Upper Bound Axiom = Completeness Axiom)** Every nonempty subset of $\mathbb{R}$ that is bounded above has a supremum (in $\mathbb{R}$).
If A is not bounded above (below), let $\sup A = \infty$ ($\inf A = -\infty$).

4. Let (a_n) be a real sequence, define

$$\liminf a_n = \sup_m \inf_{n \geq m} a_n \quad \text{and} \quad \limsup a_n = \inf_m \sup_{n \geq m} a_n.$$

Then define $\lim a_n$ to be a if $\liminf a_n = a = \limsup a_n$. In this definition, a is allowed to be $\pm\infty$.
Verify that when $a \in \mathbb{R}$, this definition is equivalent to the usual "ε-N" definition of limit.

Date: August 2017.

1.2. **Sets.**

Let Ω be a set.

1. Let A_α be a family of subsets of Ω. Then

$\cap_\alpha A_\alpha = \{\omega \in \Omega : \omega \in A_\alpha \text{ for all } \alpha\}, \quad \cup_\alpha A_\alpha = \{\omega \in \Omega : \omega \in A_\alpha \text{ for some } \alpha\}.$

2. If $A, B \subseteq \Omega$, the **difference set** $A \setminus B = \{\omega \in \Omega : \omega \in A \text{ and } \omega \notin B\}$.
The **complement** of a subset A of Ω is $A^c = \Omega \setminus A = \{\omega \in \Omega : \omega \notin A\}$.

3. DeMoivre's Theorem

$$(\cup A_\alpha)^c = \cap A_\alpha^c, \quad (\cap A_\alpha)^c = \cup A_\alpha^c.$$

4. Let A_k be a sequence of subsets of Ω. Then

$$\{\omega \in \Omega : \omega \in A_k \text{ for infinitely many } k\} = \cap_{n=1}^\infty \cup_{k=n}^\infty A_k,$$
$$\{\omega \in \Omega : \omega \in A_k \text{ for all but finitely many } k\} = \cup_{n=1}^\infty \cap_{k=n}^\infty A_k.$$

The first set is also called $\limsup_k A_k$ and the second set $\liminf_k A_k$.

1.3. **Sets and functions in $\mathbb{R}^n$.**

1. The **norm (= length)** of an element $x = (x_1, \dots, x_n)$ in $\mathbb{R}^n$ is

$$\|x\| = \sqrt{\sum_{k=1}^n |x_k|^2}.$$

2. An **open ball** in $\mathbb{R}^n$ is a set of the form

$$B(x, r) = \{y \in \mathbb{R}^n : \|y - x\| < r\}, \quad \text{where } x \in \mathbb{R}^n \text{ and } r > 0.$$

3. A set U in $\mathbb{R}^n$ is **open** if for any $x \in U$, there exists $r > 0$ so that $B(x, r) \subseteq U$. A set $V \in \mathbb{R}^n$ is **closed** if V^c is open.

4. If U_α are open sets in $\mathbb{R}^n$, then $\cup U_\alpha$ is open.
If $U_1, \dots, U_n$ (finitely many!) are open sets in $\mathbb{R}^n$, then $U_1 \cap \cdots \cap U_n$ is open.

5. Let S be a subset of $\mathbb{R}^n$. A subset U of S is **relatively open (relatively closed)** in S if there is an open set (closed set) W in $\mathbb{R}^n$ so that $U = W \cap S$.

6. Let S be a subset of $\mathbb{R}^n$. A function $f : S \to \mathbb{R}$ is **continuous** at a point $s \in S$ if for any $\varepsilon > 0$, there exists $\delta > 0$ so that $|f(x) - f(s)| < \varepsilon$ for all $x \in B(s, \delta) \cap S$.
f is **continuous on** S if it is continuous at all points in S.

7. **Theorem**. Let S be a subset of $\mathbb{R}^n$. A function $f : S \to \mathbb{R}$ is continuous on S if and only if for any open (closed) set U in $\mathbb{R}$,

$$f^{-1}(U) = \{x \in S : f(x) \in U\} \text{ is relatively open (closed) in } S.$$

1.4. **Countable sets.**

A set A is **countable** if it is either finite or there is a bijection $f : \mathbb{N} \to A$. A set that is not countable is **uncountable**. Examples of countable sets include $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$. Examples of uncountable sets are $\mathbb{R}$, $\mathbb{R} \setminus \mathbb{Q}$. More generally, any nonempty real interval is uncountable. The Cantor Set (which we will define formally in §7) is uncountable.

Proposition.
a. A subset of a countable set is countable.
b. If A_k is a sequence of countable sets, then $\cup_{k=1}^{\infty} A_k$ is countable.
c. Let $f : A \to B$ be an injective function. If B is countable, then so is A.
d. Let $f : A \to B$ be a surjective function. If B is uncountable, then so is A.

2. Review of Riemann integration

Definition. Let $[a,b]$ be a closed bounded interval in $\mathbb{R}$.
A **partition** of $[a,b]$ is a finite set $P = \{a = x_0 < x_1 < \cdots < x_n = b\}$. The **mesh** of P is $|P| = \max_{1\leq k\leq n}(x_k - x_{k-1})$.
A **pointed partition** of $[a,b]$ is an ordered pair $(P,(t_k)_{k=1}^n)$ so that $P = \{a = x_0 < x_1 < \cdots < x_n = b\}$ is a partition of $[a,b]$ and $x_{k-1} \leq t_k \leq x_k$ for all k.

Definition. Let $f : [a,b] \to \mathbb{R}$ be a bounded function on a closed bounded interval $[a,b]$ in $\mathbb{R}$.

(1) If $P = \{a = x_0 < x_1 < \cdots < x_n = b\}$ is a partition of $[a,b]$, define

$$U(f,P) = \sum_{k=1}^{n} \sup_{x\in[x_{k-1},x_k]} f(x)\cdot(x_k - x_{k-1}),$$
$$L(f,P) = \sum_{k=1}^{n} \inf_{x\in[x_{k-1},x_k]} f(x)\cdot(x_k - x_{k-1}).$$

These are called the **upper sum** and **lower sum** of f with respect to P respectively.

(2) If $(P,(t_k)_{k=1}^n)$ is a pointed partition of $[a,b]$, define

$$S(f,P,(t_k)_{k=1}^n) = \sum_{k=1}^{n} f(t_k)\cdot(x_k - x_{k-1}).$$

It is called the **Riemann sum** of f with respect to the pointed partition $(P,(t_k)_{k=1}^n)$.

Lemma 1. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function on a closed bounded interval $[a,b]$ in $\mathbb{R}$.*

(1) *Suppose that P and Q are partitions of $[a,b]$ so that $P \subseteq Q$. Then $L(f,P) \leq L(f,Q) \leq U(f,Q) \leq U(f,P)$.*
(2) *If $(P,(t_k)_{k=1}^n)$ is a pointed partition of $[a,b]$, then*

$$L(f,P) \leq S(f,P,(t_k)_{k=1}^n) \leq U(f,P).$$

Definition. Let $f : [a,b] \to \mathbb{R}$ be a bounded function on a closed bounded interval $[a,b]$ in $\mathbb{R}$.

(1) f is **Darboux integrable** to a number $I \in \mathbb{R}$ if

$$\inf_P U(f,P) = I = \sup_P L(f,P).$$

(2) f is **Riemann integrable** to a number $I \in \mathbb{R}$ if for any $\varepsilon > 0$, there exists $\delta > 0$ so that

$$|S(f,P,(t_k)_{k=1}^n) - I| < \varepsilon$$

for any pointed partition $(P,(t_k)_{k=1}^n)$ of $[a,b]$ such that $|P| < \delta$.

Theorem 2. *Let $f : [a, b] \to \mathbb{R}$ be a bounded function on a closed bounded interval $[a, b]$ in $\mathbb{R}$ and let $I \in \mathbb{R}$. Then f is Darboux integrable to I if and only if it is Riemann integrable to I.*

[Sketch. Suppose that f is Darboux integrable. Let $\varepsilon > 0$. Choose a partition P_0 such that $U(f, P_0) < I + \varepsilon$ and $L(f, P_0) > I - \varepsilon$. Suppose that $P_0 = \{a = x_0 < \cdots < x_n = b\}$ and let $M > 0$ be such that $|f(x)| \leq M$ for all $x \in [a, b]$. If P is a partition such that

$$|P| < \min\{\min_{1 \leq k \leq n}(x_k - x_{k-1}), \frac{\varepsilon}{2M(n-1)}\},$$

then $U(f, P) \leq U(f, P_0) + \varepsilon < I + 2\varepsilon$.]

The common value I in the theorem in called the **Riemann integral** of f and is denoted by $\int_a^b f$ or $\int_a^b f(x)\,dx$.

An illustration. Sum the numbers

$$\underline{3\ \ 3\ \ 7\ \ 1\ \ 3\ \ 7\ \ 2\ \ 3\ \ 7\ \ 1\ \ 7\ \ 3\ \ 2\ \ 1\ \ 3\ \ 7\ \ 1\ \ 1\ \ 2\ \ 7\ \ 1\ \ 3}.$$

3. σ-ALGEBRAS AND MEASURES

Let Ω be a set. The collection of all subsets of Ω is called the **power set** of Ω and is denoted by $\mathcal{P}(\Omega)$.
A σ-**algebra** on Ω is a set $\Sigma \subseteq \mathcal{P}(\Omega)$ so that $\emptyset \in \Sigma$, $A^c \in \Sigma$ if $A \in \Sigma$, and $\cup_{k=1}^{\infty} A_k \in \Sigma$ for any sequence $A_1, A_2, \dots$ in Σ.

Example 3.
a. Let Ω be any set. Then $\mathcal{P}(\Omega)$ is a σ-algebra on Ω.
b. Let Ω be any set and let $A \subseteq \Omega$. Then $\Sigma = \{\emptyset, A, A^c, \Omega\}$ is a σ-algebra on Ω.
c. Let Ω be any set. Let Σ consists of all subsets A of Ω so that either A or A^c is countable. Then Σ is a σ-algebra.

Proposition 4. *Let Σ be a σ-algebra on a set Ω. If $A, B \in \Sigma$, then Ω, $A \cap B$ and $A \setminus B$ are in Σ. If $A_1, A_2, \dots$ is a sequence in Σ, then $\cap_{k=1}^{\infty} A_k \in \Sigma$.*

Proposition 5. *Let $\mathcal{S} \subseteq \mathcal{P}(\Omega)$. The intersection of all σ-algebras on Ω that contains $\mathcal{S}$ is a σ-algebra on Ω, denoted by $\sigma(\mathcal{S})$. It is the smallest σ-algebra on Ω containing $\mathcal{S}$.*

Example 6.
a. Let A_α be subsets of Ω so that $\cup A_\alpha = \Omega$ and that $A_\alpha \cap A_\beta = \emptyset$ if $\alpha \neq \beta$. Denote by $\mathcal{S}$ the set of all A_α's. Then $\sigma(S)$ consists of all sets that are countable unions of A_α's or complements of countable unions of A_α's.
b. In particular, let Ω be a set and let $\mathcal{S} = \{\{\omega\} : \omega \in \Omega\}$. Then a set A is in $\sigma(\mathcal{S})$ if and only if either A is countable or A^c is countable.
c. Let $\Omega = \mathbb{R}^n$ and let $\mathcal{O}$ be the set of all open sets in $\mathbb{R}^n$. Then $\sigma(\mathcal{O})$ is called the **Borel sets***, denoted by $\mathcal{B}$ or $\mathcal{B}_n$.*
d. Let $a = (a_1, \dots, a_n), b = (b_1, \dots, b_n)$ be such that $-\infty \leq a_k \leq b_k \leq \infty$ for $1 \leq k \leq n$. The **half-open interval** *$[a, b)$ is the set*

$$[a, b) = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n : a_k \leq x_k < b_k \text{ for } 1 \leq k \leq n\}.$$

Denote by $\mathcal{H}_n$ the set of half-open intervals in $\mathbb{R}^n$. Then $\sigma(\mathcal{H}_n)$ is equal to the Borel sets.

Examples of Borel sets include the following:
(i) All closed sets;
(ii) All G_δ sets, i.e., all sets of the form $\cap_{k=1}^{\infty} U_k$, where U_k is a sequence of open sets;
(iii) All F_σ sets, i.e., all sets of the form $\cup_{k=1}^{\infty} F_k$, where F_k is a sequence of closed sets;
(ii) All $G_{\delta\sigma}$ sets, i.e., all sets of the form $\cup_{k=1}^{\infty} A_k$, where A_k is a sequence of G_δ sets;
(iii) All $F_{\sigma\delta}$ sets, i.e., all sets of the form $\cap_{k=1}^{\infty} A_k$, where A_k is a sequence of F_σ sets; etc.

Definition. Let Σ be a σ-algebra on a set Ω. A **measure** on Σ is a function $\mu : \Sigma \to [0, \infty]$ so that $\mu(\emptyset) = 0$ and that $\mu(\cup_k A_k) = \sum_{k=1}^{\infty} \mu(A_k)$ for any sequence of disjoint sets $A_k \in \Sigma$.

Proposition 7. *Let μ be a measure on a σ-algebra Σ and let $A, B \in \Sigma$ so that $A \subseteq B$. Then $\mu(A) \leq \mu(B)$. If $\mu(B) < \infty$, then $\mu(B \backslash A) = \mu(B) - \mu(A)$.*

Example 8. *Let Ω be a set. Define $\mu : \mathcal{P}(\Omega) \to [0, \infty]$ by $\mu(A) = \#A$ if A is a finite set and $\mu(A) = \infty$ if A is an infinite set. Then μ is a measure on $\mathcal{P}(\Omega)$. It is called the* **counting measure** *on Ω.*

A **measurable space** is a pair (Ω, Σ), where Ω is a set and Σ is a σ-algebra on Ω. A **measure space** is a triple (Ω, Σ, μ), where (Ω, Σ) is a measurable space and $\mu : \Sigma \to [0, \infty)$ is a measure on Σ. If (Ω, Σ, μ) is a measure space, a set $A \in \Sigma$ is said to be **measurable** or Σ**-measurable** or μ**-measurable**.

Proposition 9. *Let (Ω, Σ, μ) be a measure space and let A_k be a sequence of sets in Σ.*

(1) *If $A_1 \subseteq A_2 \subseteq \cdots$ and $A = \cup_{k=1}^{\infty} A_k$, then $\mu(A) = \lim_{k\to\infty} \mu(A_k)$.*
(2) *Assume that $A_1 \supseteq A_2 \supseteq \cdots$ and $A = \cap_{k=1}^{\infty} A_k$. If $\mu(A_k) < \infty$ for some k, then $\mu(A) = \lim_{k\to\infty} \mu(A_k)$.*

Let (Ω, Σ, μ) be a measure space. A set $N \in \Sigma$ is a μ**-null set** if $\mu(N) = 0$. Let $P(\omega)$ be a property that each $\omega \in \Omega$ may or may not possess. We say the $P(\omega)$ holds **almost everywhere** (a.e.) if there is a μ-null set N so that $P(\omega)$ holds for all $\omega \notin N$.

4. Measurable functions

Let (Ω, Σ) be a given measurable space.

Definition. A function $f : \Omega \to [-\infty, \infty]$ is **measurable** if the set

$$\{f > a\} = \{\omega \in \Omega : f(\omega) > a\} \in \Sigma$$

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

Proposition 10. *Let $f : \Omega \to [-\infty, \infty]$ be measurable.*

(1) *For any $a \in \mathbb{R}$, the sets*

$$\{f < a\} = \{\omega \in \Omega : f(\omega) < a\},\ \{f = a\} = \{\omega \in \Omega : f(\omega) = a\},$$
$$\{f \geq a\} = \{\omega \in \Omega : f(\omega) \geq a\},\ \{f \leq a\} = \{\omega \in \Omega : f(\omega) \leq a\}$$

all belong to Σ.

(2) *The sets*

$$\{f = \infty\} = \{\omega \in \Omega : f(\omega) = \infty\} \text{ and } \{f = -\infty\} = \{\omega \in \Omega : f(\omega) = -\infty\}$$

belong to Σ.

Suppose that $a, b \in [-\infty, \infty]$. Define $a + b$ in the obvious way except if $\{a = \infty \text{ and } b = -\infty\}$ or $\{a = -\infty \text{ and } b = \infty\}$. Similarly, ab is defined in the obvious way except that we take $0 \cdot \pm\infty = \pm\infty \cdot 0 = 0$. Then $a - b$ is defined to be $a + (-1)b$. If $f, g : \Omega \to [-\infty, \infty]$, define $f + g : \Omega \to [-\infty, \infty]$ by

$$(f + g)(\omega) = \begin{cases} f(\omega) + g(\omega) & \text{if the sum is defined;} \\ 0 & \text{if the sum is undefined.} \end{cases}$$

Proposition 11. *Let f, g be Σ-measurable functions and let c be a real number.*

(1) *Then the functions cf, $|f|$, fg, $f + g$, $f - g$, $\max\{f, g\}$, $\min\{f, g\}$ are measurable.*
(2) *If (f_n) is a sequence of measurable functions, then $\inf f_n$, $\sup f_n$, $\liminf f_n$ and $\limsup f_n$ are measurable functions.*
(3) *If (f_n) is a sequence of measurable functions and $f(\omega) = \lim_n f_n(\omega)$ exists (in $[-\infty, \infty]$), then f is measurable.*
(4) *If $h : [-\infty, \infty] \to [-\infty, \infty]$ is a continuous function, then $h \circ f$ is measurable.*

For any $f : \Omega \to [-\infty, \infty]$, let $f^+ = \max\{f, 0\}$ and $f^- = -\min\{f, 0\}$.

Proposition 12. *For any $f \in \Omega \to [-\infty, \infty]$, f^+ and f^- are nonnegative functions such that $f = f^+ - f^-$. f is measurable if and only if f^+ and f^- are both measurable.*

A measure space (Ω, Σ, μ) is said to be **complete** if any subset of a μ-null set belongs to Σ.

Proposition 13. *Let (Ω, Σ, μ) be a complete measure space. If f is measurable and $g = f$ a.e., then g is measurable.*

Example 14. *Let $\Omega = \mathbb{R}$ and define $\Sigma \subseteq \mathcal{P}(\Omega)$ by $A \in \Sigma$ if and only if either A or A^c is a countable set (cf. Example 3c). Then (Ω, Σ) is a measurable space. A function $f : \Omega \to [-\infty, \infty]$ is measurable if and only if $\{f = a\} \in \Sigma$ for each $a \in [-\infty, \infty]$ and f is countably-valued, i.e., $f(\Omega)$ is a countable set.*

Definition. Let $A \subseteq \Omega$. The **characteristic function** of A is the function $\chi_A : \Omega \to \mathbb{R}$ so that $\chi_A(\omega) = 1$ if $\omega \in A$ and 0 otherwise.
A **step function** is a function of the form $\sum_{k=1}^{n} c_k \chi_{A_k}$, where $n \in \mathbb{N}$, $c_k \in \mathbb{R}$ and $A_k \in \Sigma$ for all k.

Proposition 15.
(1) *A function f on Ω is a step function if and only if it is measurable and $f(\Omega)$ is a finite subset of $\mathbb{R}$.*
(2) *Let f and g be step functions and let $c \in \mathbb{R}$. Then cf and $f + g$ are step functions.*
(3) *If f is a step function, then it can be written in the form $\sum_{k=1}^{m} a_k \chi_{B_k}$, where $a_k \in \mathbb{R}$, $B_k \in \Sigma$ and $B_k \cap B_j = \emptyset$ if $k \neq j$.*

Theorem 16. *A function $f : \Omega \to [-\infty, \infty]$ is measurable if and only if there is a sequence of step functions f_n so that $\lim f_n(\omega) = f(\omega)$ for all $\omega \in \Omega$.*
If $f \geq 0$, then f_n can be chosen so that $0 \leq f_n \leq f_{n+1}$ for all n.
If f is bounded, then f_n can be chosen to converge to f uniformly on Ω.

5. Lebesgue integration

In this section, let (Ω, Σ, μ) be a given measure space.

Recall from Proposition 15 that a function f defined on Ω is a step function if and only if it is measurable and takes only finitely many real values. Denote the set of step functions by $S(\Omega, \Sigma, \mu)$ and the set of nonnegative functions in $S(\Omega, \Sigma, \mu)$ by $S_+(\Omega, \Sigma, \mu)$. The definition of the Lebesgue integral begins with an elementary integral on $S_+(\Omega, \Sigma, \mu)$, which gets extended step-by-step.

Proposition 17. *Let $f \in S_+(\Omega, \Sigma, \mu)$. Define $I(f) = \sum a_k \mu\{f = a_k\}$, where the sum is taken over all a_k's in the range of f (recall that this is a finite set).*

1. *Suppose that $f, g \in S_+(\Omega, \Sigma, \mu)$ and $c \geq 0$. Then $I(cf + g) = cI(f) + I(g)$.*
2. *(Monotonicty) $I(f) \geq I(g)$ if $f, g \in S_+(\Omega, \Sigma, \mu)$ and $f \geq g$.*
3. *$I(f) = I(g)$ if $f, g \in S_+(\Omega, \Sigma, \mu)$ and $f = g$ a.e.*

[Sketch. The key idea is that if $f = \sum_{k=1}^m b_k \chi_{B_k}$, where $b_k \geq 0$, $B_k \in \Sigma$, and $B_j \cap B_k = \emptyset$ if $j \neq k$, then $I(f) = \sum_{k=1}^m b_k \mu(B_k)$.]

Definition. Let f be a nonnegative measurable function. Define

$$\int f = \int f\, d\mu = \int_\Omega f\, d\mu = \sup\{I(g) : g \in S_+(\Omega, \Sigma, \mu), g \leq f\}.$$

Lemma 18. *Let (h_n) be a nondecreasing sequence in $S_+(\Omega, \Sigma, \mu)$ and let $h \in S_+(\Omega, \Sigma, \mu)$. Suppose that $\lim_n h_n(\omega) \geq h(\omega)$ for all $\omega \in \Omega$. Then $\lim_n I(h_n) \geq I(h)$.*

Proposition 19. *Let f be a nonnegative measurable function.*

1. *$\int f \geq 0$ and $\int f = 0$ if and only if $f = 0$ a.e.*
2. *$\int f\, d\mu = I(f)$ if f is a nonnegative step function.*
3. *$\int cf\, d\mu = c \int f\, d\mu$ for any $c \geq 0$.*
4. *Suppose that g is a nonnegative measurable function. Then*
$$\int (f+g)\, d\mu = \int f\, d\mu + \int g\, d\mu.$$
5. *Suppose that g is a nonnegative measurable function such that $f \geq g$. Then $\int f\, d\mu \geq \int g\, d\mu$.*
6. *(B. Levi's Theorem = Monotone Convergence Theorem) Let (f_n) be a sequence of nonnegative measurable functions so that $f_n \leq f_{n+1}$ for all n. Set $f = \lim_n f_n$. Then $\int f\, d\mu = \lim_n \int f_n\, d\mu$.*

Definition of Lebesgue integral. Let $f : \Omega \to [-\infty, \infty]$ be a measurable function. Define

$$\int f\, d\mu = \int f^+\, d\mu - \int f^-\, d\mu$$

if the right hand side is defined (i.e., not of the form $\infty - \infty$). Otherwise, say that the integral on the left is undefined. f is said to be **integrable** or **μ-integrable** if $\int f^+ \, d\mu$ and $\int f^- \, d\mu$ are both finite.

Proposition 20. *Let f be a measurable function.*

1. *$\int |f| = 0$ if and only if $f = 0$ a.e.*
2. *Suppose that g is a measurable function and $c \in \mathbb{R}$. Then*
$$\int (cf + g) \, d\mu = c \int f \, d\mu + \int g \, d\mu$$
provided both sides are defined.
3. *Suppose that g is a measurable function such that $f \geq g$. Then $\int f \, d\mu \geq \int g \, d\mu$ if both integrals are defined.*
4. *f is integrable if and only if $|f|$ is integrable.*

Proposition 21. *(Chebyshev's inequality) Let f be a measurable function. For any $0 < a \in \mathbb{R}$,*
$$a\mu\{|f| \geq a\} \leq \int |f| \, d\mu.$$

A set of the form $\cup_{k=1}^{\infty} A_k$, where $A_k \in \Sigma$ and $\mu(A_k) < \infty$ for all k, is said to be **σ-finite**.

Corollary 22. *Let f be an integrable function. Then the set $\{f \neq 0\}$ is σ-finite and the set $\{|f| = \infty\}$ is μ-null.*

Example 23. *Consider the counting measure space $(\Omega, \mathcal{P}(\Omega), \mu)$. Every function is measurable.*
If f is a nonnegative (measurable) function on Ω, then
$$\int f \, d\mu = \sup \sum_{\omega \in F} f(\omega),$$
where the sup is taken over all finite sets $F \subseteq \Omega$.
If f is a (measurable) function on Ω, then f is integrable if and only if $\sup \sum_{\omega \in F} |f(\omega)| < \infty$, where the sup is taken over all finite sets $F \subseteq \Omega$.
If f is integrable and $\int f \, d\mu = a \in \mathbb{R}$, then for any $\varepsilon > 0$, there exists a finite set $F \subseteq \Omega$ so that
$$|\sum_{\omega \in G} f(\omega) - a| < \varepsilon$$
for any finite set $G \subseteq \Omega$ such that $F \subseteq G$.

6. Construction of measures

So far we have described how to define the Lebesgue integral *given a measure space*. However, examples of measure spaces are rather lacking except for the counting measure. In particular, we do not have a measure space that would allow us to do "normal" integration on the real line. In this section, we will describe a general procedure to construct measures. In particular, the procedure yields **the** Lebesgue measure on $\mathbb{R}$ (even $\mathbb{R}^n$). Integration with respect to Lebesgue measure generalizes Riemann integration.

We begin with a nonempty set Ω.

Definition. An **outer measure** on Ω is a function $\mu^* : \mathcal{P}(\Omega) \to [0, \infty]$ so that

(1) $\mu^*(\emptyset) = 0$.
(2) (Monotonicity) $\mu^*(A) \leq \mu^*(B)$ if $A \subseteq B$.
(3) (Countable subadditivity) $\mu^*(\cup A_k) \leq \sum \mu^*(A_k)$ for all $A_k \subseteq \Omega$.

Outer measures can be generated rather easily.

Proposition 24. *Suppose that $\emptyset \in \mathcal{S} \subseteq \mathcal{P}(\Omega)$. Let $\mu_0 : \mathcal{S} \to [0, \infty]$ be any function such that $\mu_0(\emptyset) = 0$. Define $\mu^* : \mathcal{P}(\Omega) \to [0, \infty]$ by*

$$\mu^*(A) = \inf\{\sum_{n=1}^{\infty} \mu_0(A_n) : A_n \in \mathcal{S}, A \subseteq \cup_{n=1}^{\infty} A_n\},$$

where the infimum is taken to be ∞ if the set on the right is empty. Then μ^* is an outer measure on Ω.

Theorem 25. *(Carathéodory's definition of measurable sets) Let μ^* be an outer measure on Ω. Let Σ be the collection of subsets of Ω so that $E \in \Sigma$ if and only if*

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \setminus E) \text{ for all } A \subseteq \Omega. \tag{1}$$

Then Σ is a σ-algebra. Define $\mu : \Sigma \to [0, \infty]$ by $\mu(A) = \mu^*(A)$ for all $A \in \Sigma$. Then μ is a measure on Σ.

Remark. In equation (1),

$$\mu^*(A) \leq \mu^*(A \cap E) + \mu^*(A \setminus E)$$

always holds by subadditivity of μ^*. Thus (1) holds if and only if the reverse inequality holds. Sets in Σ are said to be μ^* or μ-measurable.

[Sketch. Clearly $E \in \Sigma \iff E^c \in \Sigma$. A Venn diagram helps to show that $E_1, E_2 \in \Sigma \implies E_1 \cap E_2 \in \Sigma$. Then show that if (E_k) is a disjoint sequence of sets in Σ, then $\cup_k E_k \in \Sigma$.]

In summary, Carathéodory's procedure is as follows:

$$\begin{aligned} \text{Function } \mu_0 \text{ on } \mathcal{S} \subseteq \mathcal{P}(\Omega) &\to \text{Outer measure } \mu^* \text{ on } \mathcal{P}(\Omega) \\ &\to \text{measure } \mu \text{ on } \Sigma. \end{aligned}$$

Recall that a measure space (Ω, Σ, μ) is complete if any subset of a μ-null set is measurable.

Proposition 26. *Let μ^* be an outer measure on Ω and let (Ω, Σ, μ) be the measure space obtained from μ^* by applying Carathédory's construction. Then (Ω, Σ, μ) is a complete measure space.*

7. Lebesgue measure on $\mathbb{R}^n$

The most important application of Carathéodory's procedure is the following.

Example 27. *Recall from Example 6d that $\mathcal{H}_n$ refers to the set of all half-open intervals in $\mathbb{R}^n$. Define $\lambda_0 : \mathcal{H}_n \to [0, \infty]$ by*

$$\lambda_0(a, b] = \prod_{k=1}^{n} (b_k - a_k)$$

for any half-open interval $(a, b]$, where $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$. By Proposition 24,

$$\lambda^*(A) = \inf\{\sum_{k=1}^{\infty} \lambda_0(A_k) : A \subseteq \cup_{k=1}^{\infty} A_k,\ A_k\text{'s are half-open inervals}\}$$

defines an outer measure on $\mathcal{P}(\mathbb{R}^n)$. We call λ^* **Lebesgue outer measure**. *The λ^*-measurable sets are called* **Lebesgue measurable sets**. *Let Σ be the set of Lebesgue measurable sets. The corresponding measure λ (which is just the restriction of λ^* onto Σ) is called* **Lebesgue measure** *on $\mathbb{R}^n$. The Lebesgue measure space $(\mathbb{R}^n, \Sigma, \lambda)$ is complete.*

Proposition 28. *All Borel sets in $\mathbb{R}^n$ are Lebesgue measurable. $\lambda_0(I) = \lambda^*(I)$ for all $I \in \mathcal{H}$.*

Theorem 29. *(Regularity of Lebesgue measure) A set $A \subseteq \mathbb{R}^n$ is Lebesgue measurable if and only if for any $\varepsilon > 0$, there exists an open set O so that $A \subseteq O$ and that $\lambda^*(O \setminus A) < \varepsilon$. If A is Lebesgue measurable and $\lambda(A) < \infty$, then for any $\varepsilon > 0$, there exists a compact set $K \subseteq A$ so that $\lambda(A \setminus K) < \varepsilon$.*

Let E be a Lebesgue measurable subset of $\mathbb{R}^n$. The set of all Lebesgue measurable subsets of E is a σ-algebra (denoted Σ_E) on E. The restriction of Lebesgue measure on $\mathbb{R}^n$ to this σ-algebra is a measure. In this way, we obtain a measure space (E, Σ_E, λ), referred to as the Lebesgue measure space on E.

Theorem 30. *(Egorov's Theorem) Let (Ω, Σ, μ) be a measure space so that $\mu(\Omega) < \infty$. Assume that $f_n : \Omega \to \mathbb{R}$ is a sequence of μ-measurable functions that converges almost everywhere to a μ-measurable function $f : \Omega \to \mathbb{R}$. For any $\varepsilon > 0$, there exist a set $K \in \Sigma$ so that $\mu(K^c) < \varepsilon$ and that $f_n \to f$ uniformly on K. Moreover, if $(\Omega, \Sigma, \mu) = (E, \Sigma_E, \lambda)$, where E is a Lebesgue measurable set in $\mathbb{R}^n$ such that $\lambda(E) < \infty$, then the set K can be chosen to be compact.*

[Sketch. Let $E_{mn} = \{|f - f_n| > \frac{1}{m}\}$. For each m, $\cap_{k=1}^{\infty} \cup_{n=k}^{\infty} E_{mn}$ is a null set. Choose $k_1 < k_2 < \cdots$ so that $\mu(F_m) = \mu(\cup_{n=k_m}^{\infty} E_{mn}) < \frac{\varepsilon}{2^m}$. Take $K = (\cup_{m=1}^{\infty} F_m)^c$.]

Theorem 31. *(Lusin's Theorem) Let E be a Lebesgue measurable subset of $\mathbb{R}^n$ such that $\lambda(E) < \infty$. A function $f : E \to \mathbb{R}$ is Lebesgue measurable* *if and only if for any $\varepsilon > 0$, there is a compact set $K \subseteq E$ such that $\lambda(E \setminus K) < \varepsilon$ and that $f_{|K}$ is a continuous function on K.*

[Sketch. First prove it for step functions. If f is a bounded Lebesgue measurable function, there is a sequence of step functions that converges to f uniformly. The general case can be reduced to the case for bounded functions.]

Collectively, Theorems 29, 30 and 31 are known as "Littlewood's Three Principles". They say that, roughly speaking, Lebesgue measurable sets are almost open, convergence a.e. is almost uniform convergence and measurability with respect to Lebesgue measure is almost continuity.

Definition. Let A be a subset of $\mathbb{R}^n$, $x \in \mathbb{R}^n$ and $0 \neq c \in \mathbb{R}$. Define

$$x + A = \{x + \omega : \omega \in A\} \quad \text{and} \quad cA = \{c\omega : \omega \in A\}.$$

Proposition 32. *Let A be a subset of $\mathbb{R}^n$, $x \in \mathbb{R}^n$, $0 \neq c \in \mathbb{R}$ and let $f : \mathbb{R}^n \to [-\infty, \infty]$ be a function.*

(1) $\lambda^*(x + A) = \lambda^*(A)$ *and* $\lambda^*(cA) = |c|^n \lambda^*(A)$.
(2) *A is Lebesgue measurable if and only if $x + A$ is Lebesgue measurable, if and only if cA is Lebesgue measurable.*
(3) *Define g and h by $g(\omega) = f(x + \omega)$ and $h(\omega) = f(\frac{\omega}{c})$ respectively. Then f is measurable if and only if g is measurable if and only if h is measurable. f is integrable if and only if g is integrable if and only if h is integrable, in which case*

$$\int g \, d\lambda = \int f \, d\lambda \quad \textit{and} \quad \int h \, d\lambda = |c|^n \int f \, d\lambda.$$

Theorem 33. *Let $f : [a, b] \to \mathbb{R}$ be a bounded function on a closed bounded interval in $\mathbb{R}$. Then f is Riemann integrable if and only if f is continuous a.e. on $[a, b]$. If f is Riemann integrable on $[a, b]$, then f is Lebesgue integrable on $[a, b]$ and*

$$\int_{[a,b]} f \, d\lambda = \textit{Riemann} - \int_a^b f(x) \, dx.$$

Sketch. If f is Riemann integrable and hence Darboux integrable to I, choose partitions P_n so that $P_n \subseteq P_{n+1}$ and that $L(f, P_n) \uparrow I$ and $U(f, P_n) \downarrow I$. We may assume that $U(f, P_n) - L(f, P_n) < \frac{1}{4^n}$. Then the total lengths of the intervals J determined by the partition P_n where $\sup_{x \in J} f(x) - \inf_{x \in J} f(x) > \frac{1}{2^n}$ is at most $\frac{1}{2^n}$. Let the union of these intervals be T_n. The set $\limsup T_n$ is a Lebesgue null set and f is continuous on its complement.
Conversely, suppose that the set N of discontinuities f is a null set. Construct a cover of $[a, b]$ by finitely many intervals $J_1, \dots, J_m$, $L_1, \dots, L_n$ so that $|f(x) - f(y)|$ if small if $x, y \in J_i$ for some i, and $\sum_{i=1}^n \lambda(L_i)$ is small. Determine a partition P of $[a, b]$ by taking the end points of all J_i and L_i (and the points a, b.) Show that $U(f, P) - L(f, P)$ is small.

Theorem 34. *Let A be a subset of $\mathbb{R}$ with $\lambda^*(A) > 0$. There exists a subset B of A that is not Lebesgue measurable.*

Sketch. We may assume that A is bounded. In fact, assume that $A \subseteq [-1, 1]$. Define an equivalence relation on A by $a \sim b$ if and only if $a - b \in \mathbb{Q}$. Let E be a subset of A that consists of exactly one element from each equivalence class. Then

$$A \subseteq \cup_{q\in\mathbb{Q}\cap[-2,2]}(q + E) \subseteq [-3, 3].$$

In particular, $\lambda^*(E) > 0$. Also, $(q + E) \cap (q' + E) = \emptyset$ if q, q' are distinct numbers in $\mathbb{Q}$. If E is measurable, it would follow that $\lambda([-3, 3]) = \infty$.

Example 35. *(Cantor set) Let $J_0 = [0, 1]$. If J_k is defined and is a union of finitely many closed intervals, let J_{k+1} be the set obtained from J_k by removing the open middle third from each of the intervals in J_k. Precisely, at stage k, suppose that J_k has been constructed and is a union of 2^k pairwise disjoint closed bounded intervals. Label the open middle thirds of these intervals as $I_{2^k}, I_{2^k+1}, \ldots, I_{2^{k+1}-1}$, arranged from left to right. Set $J_{k+1} = J_k \setminus \cup_{j=0}^{2^k-1} I_{2^k+j}$. Thus, e.g.,*

$$J_1 = [0, 1/3] \cup [2/3, 1],$$
$$J_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1].$$

The **Cantor set** *is the set $C = \cap_{k=0}^{\infty} J_k$. The Cantor set is a closed set and hence Lebesgue measurable. It has Lebesgue measure 0.*

Example 36. *(Cantor-Lebesgue function) Retain the notation of the previous example. Define $f_k : [0, 1] \to \mathbb{R}$ to be the continuous function which satisfies*

$$f_k(0) = 0, \quad f_k(1) = 1,$$
$$f_k(x) = \frac{2j+1}{2^{m+1}} \quad \text{if} \quad x \in I_{2^m+j}, 0 \le m < k, 0 \le j < 2^m,$$
$$f_k \text{ is linear on each interval of } J_k.$$

For example, f_1 and f_2 are obtained by linearly interpolating the following values:

$$f_1(0) = 0, f_1(1) = 1, f_1 = 1/2 \text{ on } [1/3, 2/3]$$
$$f_2(0) = 0, f_2(1) = 1, f_2 = \begin{cases} 1/4 & \text{on } [1/9, 2/9], \\ 1/2 & \text{on } [1/3, 2/3], \\ 3/4 & \text{on } [7/9, 8/9]. \end{cases}$$

The sequence $(f_k)_{k=1}^{\infty}$ converges uniformly to a function f on $[0, 1]$, called the **Cantor-Lebesgue function**. *f is continuous and nondecreasing on $[0, 1]$ and maps the Cantor set C onto $[0, 1]$. In particular, C is uncountable.*

8. Convergence theorems

A big advantage of the Lebesgue integral over the Riemann integral is the convergence theorems, i.e., the ability to switch limits with integration under relatively general hypotheses. Let (Ω, Σ, μ) be a measure space.

Proposition 37. *(Fatou's Lemma) Let (f_n) be a sequence of nonnegative measurable functions. Then*

$$\int \liminf f_n \, d\mu \leq \liminf \int f_n \, d\mu.$$

[Sketch. Apply the Monotone Convergence Theorem to the sequence $g_n = \inf_{m \geq n} f_m$.]

Theorem 38. *(Lebesgue Dominated Convergence Theorem) Let (f_n) be a sequence of measurable functions and let f be a measurable function so that $\lim f_n(\omega) = f(\omega)$ for almost all $\omega \in \Omega$. Assume that there is an integrable function g so that $|f_n| \leq g$ for all n. Then*

$$\lim \int f_n \, d\mu = \int f \, d\mu.$$

[Sketch. Apply Fatou's Lemma to $f_n + g$ and $g - f_n$.]

Corollary 39. *(Bounded Convergence Theorem) Let (Ω, Σ, μ) be a finite measure space, i.e., $\mu(\Omega) < \infty$. If (f_n) is a bounded sequence of measurable functions (i.e., there exists $M \in \mathbb{R}$ so that $|f_n(\omega)| \leq M$ for all n and all ω) and there is a measurable function f so that $\lim f_n(\omega) = f(\omega)$ for almost all $\omega \in \Omega$, then*

$$\lim \int f_n \, d\mu = \int f \, d\mu.$$

Example 40. *Let $f_n = n\chi_{(0,\frac{1}{n})}$ on $\mathbb{R}$ with Lebesgue measure. This example shows that equality fails to hold in Fatou's Lemma in general, and that exchange of limit with integration fails in general.*

The following is a well-known application of LDCT.

Example 41. *Let $f = f(x, y) : \mathbb{R}^2 \to \mathbb{R}$ be a function whose partial derivative with respect to y exists for all $(x, y) \in \mathbb{R}^2$. Assume that $f(x, y)$ and $\frac{\partial f}{\partial y}(x, y)$ are Lebesgue integral functions of x for each $y \in \mathbb{R}$. Also suppose that there is a Lebesgue integrable function g on $\mathbb{R}$ so that $|\frac{\partial f}{\partial y}(x, y)| \leq g(x)$ for all x, y. Then*

$$\frac{\partial}{\partial y} \int_{\mathbb{R}} f(x, y) \, d\lambda(x) = \int_{\mathbb{R}} \frac{\partial f}{\partial y}(x, y) \, d\lambda(x).$$

[Note. Differentiation involves taking limits over real numbers. In order to use LDCT, we need to consider sequences of real numbers converging to a limit.]

We close this section with a discussion of another notion of convergence of measurable functions.

Definition. Let (f_n) be a sequence of measurable functions on a measure space (Ω, Σ, μ). We say that (f_n) **converges in measure** to a measurable function f if for any $\varepsilon > 0$, $\lim_n \mu\{|f_n - f| > \varepsilon\} = 0$.

Remark. If (f_n) converges in measure to a measurable function f and g is a measurable function, then (f_n) converges to g in measure if and only if $g = f$ a.e.

Example 42.
(a) On $\mathbb{R}$, let $f_n = \chi_{(n,n+1)}$ for all n. Then (f_n) is a sequence of Lebesgue measurable function that converges pointwise to 0 but does not converge in measure (to any measurable function).
(b) On $[0,1]$, let $f_{2^k+j} = \chi_{[\frac{j}{2^k},\frac{j+1}{2^k})}$, $j = 0,1,2,\ldots,2^k-1$, $k = 0,1,2,\ldots$. Then (f_n) converges to 0 in measure but does not converge a.e. to any measurable function.

Proposition 43. *Let (Ω, Σ, μ) be a measure space and let (f_n) be a sequence of μ-measurable functions on Ω.*

1. *If $\mu(\Omega) < \infty$ and if (f_n) converges to a <u>finite valued</u> measurable function f a.e., then (f_n) converges in measure to f.*
2. *If (f_n) converges in measure to a measurable function f, then there is a subsequence (f_{n_k}) that converges to f a.e.*

[Sketch. (1) Apply Egorov's Theorem.
(2). Choose n_k so that $\mu\{|f_{n_k} - f| > \frac{1}{2^k}\} < \frac{1}{2^k}$. The lim sup of these sets is a μ-null set and (f_{n_k}) converges to f on its complement.]

Proposition 44. *(Extended Lebesgue Dominated Convergence Theorem) Let (f_n) be a sequence of measurable functions and let f be a measurable function so that (f_n) converges to f in measure. Assume that there is an integrable function g so that $|f_n| \le g$ for all n. Then*

$$\lim \int f_n \, d\mu = \int f \, d\mu.$$

Note. Finiteness of the measure space (Ω, Σ, μ) is <u>not</u> assumed in Proposition 44.

9. L^p SPACES

Let (Ω, Σ, μ) be a measure space.

Definition. If $1 \leq p < \infty$, the set $L^p(\Omega, \Sigma, \mu)$ consists of all measurable functions $f : \Omega \to [-\infty, \infty]$ such that $\int |f|^p \, d\mu < \infty$. The set $L^\infty(\Omega, \Sigma, \mu)$ consists of all measurable function $f : \Omega \to [-\infty, \infty]$ such that

$$\inf\{M \in [-\infty, \infty] : |f| \leq M \text{ a.e.}\} \text{ is finite.} \tag{2}$$

If $f \in L^p(\Omega, \Sigma, \mu)$, $1 \leq p < \infty$, set $\|f\|_p = (\int |f|^p \, d\mu)^{1/p}$.
If $f \in L^\infty(\Omega, \Sigma, \mu)$, then set $\|f\|_\infty$ to be the infimum in (2).

Proposition 45. *Suppose that $f, g \in L^p(\Omega, \Sigma, \mu)$ and $c \in \mathbb{R}$. Then cf and $f + g$ are in $L^p(\Omega, \Sigma, \mu)$.*

[For $f + g$, use the fact that $|a + b|^p \leq 2^p(|a|^p + |b|^p)$.]

Proposition 46. *(Young's inequality) Suppose that $1 < p < \infty$ and $q = \frac{p}{p-1}$. If $a, b \geq 0$, then*

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

Proposition 47. *(Hölder's inequality) Suppose that $1 < p < \infty$ and $q = \frac{p}{p-1}$ or $p = 1$ and $q = \infty$. If $f \in L^p(\Omega, \Sigma, \mu)$ and $g \in L^q(\Omega, \Sigma, \mu)$, then $fg \in L^1(\Omega, \Sigma, \mu)$ and*

$$\int |fg| \, d\mu \leq \|f\|_p \, \|g\|_q.$$

More precisely,

$$\|f\|_p = \sup_{\|g\|_q \leq 1} |\int fg \, d\mu|.$$

[To prove Hölder's inequality, apply Young's inequality with

$$a = \frac{|f(\omega)|}{\|f\|_p} \text{ and } b = \frac{|g(\omega)|}{\|g\|_q}$$

and integrate.
If $1 < p < \infty$, the last equality can be obtained by taking g to be $\frac{|f|^{p-1}}{\||f|^{p-1}\|_q}$.]

Remark. The final equation in Proposition 47 does not always hold if $p = \infty$ and $q = 1$.

Proposition 48. *(Minkowski's inequality) Suppose that $1 \leq p \leq \infty$. If $f, g \in L^p(\Omega, \Sigma, \mu)$, then*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

[Sketch.

$$\|f+g\|_p^p \leq \int |f| \cdot |f+g|^{p-1}\, d\mu + \int |g| \cdot |f+g|^{p-1}\, d\mu.$$

Apply Hölder's inequality to the two terms on the right.]

A sequence (f_k) in $L^p(\Omega, \Sigma, \mu)$ is said to be p-**Cauchy** if for any $\varepsilon > 0$, there exists N such that $\|f_k - f_j\|_p < \varepsilon$ for all $k, j \geq N$.

Theorem 49. *(Riesz-Fischer Theorem) Suppose that* $1 \leq p \leq \infty$ *and that* (f_k) *is a p-Cauchy sequence in* $L^p(\Omega, \Sigma, \mu)$. *There is a function* $f \in L^p(\Omega, \Sigma, \mu)$ *so that* $\lim_k \|f_k - f\|_p = 0$.

[Sketch. The main step is to produce the limit function f from the p-Cauchy sequence (f_k). Choose a subsequence (f_{k_i}) so that $\|f_{k_i} - f_{k_{i+1}}\|_p \leq \frac{1}{2^i}$ for all i. Use Minkowski's inequality and the Monotone Convergence Thereom to show that the function $g = |f_{k_1}| + \sum_{i=1}^{\infty} |f_{k_i} - f_{k_{i+1}}|$ belongs to L^p. Hence the sum $f_{k_1} + \sum_{i=1}^{\infty}(f_{k_{i+1}} - f_{k_i})$ converges a.e. to a function $f \in L^p$. Use LDCT to show that $\|f_{k_i} - f\|_p \to 0$.]

10. Product measure

Let $(\Omega_1, \Sigma_1, \mu_1)$ and $(\Omega_2, \Sigma_2, \mu_2)$ be measure spaces.

Definition. A **measurable rectangle** on $\Omega_1 \times \Omega_2$ is a set of the form $A_1 \times A_2$, where $A_1 \in \Sigma_1$ and $A_2 \in \Sigma_2$. The set of all measurable rectangles is denoted by $\mathcal{R}$.

Define $\mu_0 : \mathcal{R} \to [0, \infty]$ by

$$\mu_0(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2).$$

We now apply Carathéody's procedure to μ_0 to obtain a measure space $(\Omega_1 \times \Omega_2, \Sigma, \mu)$. This is called the **product measure space** of $(\Omega_1, \Sigma_1, \mu_1)$ and $(\Omega_2, \Sigma_2, \mu_2)$. μ is called the **product measure** of μ_1 and μ_2, sometimes written as $\mu = \mu_1 \times \mu_2$. Specifically, let

$$\mu^*(A) = \inf\{\sum_k \mu_0(R_k) : R_k \in \mathcal{R}, A \subseteq \cup R_k\}$$

for all $A \subseteq \Omega_1 \times \Omega_2$. Then $E \in \Sigma$ if and only if

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \setminus E) \text{ for all } A \subseteq \Omega_1 \times \Omega_2.$$

μ is the restriction of μ^* to Σ.

Proposition 50. *Every measurable rectangle is μ-measurable, i.e., $\mathcal{R} \subseteq \Sigma$. Moreover, $\mu(R) = \mu_0(R)$ for all $R \in \mathcal{R}$.*

[Sketch. Apply the definitions. Observe that if $R = A_1 \times A_2$ and $R' = B_1 \times B_2$ are measurable rectangles, then

$R \cap R' = (A_1 \cap B_1) \times (A_2 \cap B_2)$ and $R' \setminus R = [B_1 \times (B_2 \setminus A_2)] \cup [(B_1 \setminus A_1) \times (A_2 \cap B_2)]$.]

Since Σ is a σ-algebra and $\mathcal{R} \subseteq \Sigma$, $\sigma(\mathcal{R}) \subseteq \Sigma$. In general, Σ may be strictly larger than $\sigma(\mathcal{R})$.

Example 51. *Let λ_m be Lebesgue measure on $\mathbb{R}^m$ and λ_n be Lebesgue measure on $\mathbb{R}^n$. Then the product measure $\lambda_m \times \lambda_n$ on $\mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^{m+n}$ is Lebesgue measure on $\mathbb{R}^{m+n}$.*

[Sketch. It is enough to show that the outer measure $(\lambda_m \times \lambda_n)^*$ coincides with the outer measure $(\lambda_m \times \lambda_n)^*$.]

We now discuss the main results concerning the computation of "double integral" by" iterated integration". For the rest of the section, let $(\Omega_1, \Sigma_1, \mu_1)$ and $(\Omega_2, \Sigma_2, \mu_2)$ be complete measure spaces and let $(\Omega_1 \times \Omega_2, \Sigma, \mu)$ be their product measure space.

Given a function $f : \Omega_1 \times \Omega_2 \to [-\infty, \infty]$, for any $\omega_1 \in \Omega_1$, define $f_{\omega_1} : \Omega_2 \to [-\infty, \infty]$ by $f_{\omega_1}(\omega_2) = f(\omega_1, \omega_2)$. For any $\omega_2 \in \Omega_2$, define the function $f_{\omega_2} : \Omega_1 \to [-\infty, \infty]$ similarly.

Definition. Let $f : \Omega_1 \times \Omega_2 \to [-\infty, \infty]$ be a μ-integrable function. We say that f **has property (F)**, written as $f \in (F)$, if, for $i, j \in \{1, 2\}$, $i \neq j$,

(1) For μ_i-almost all $\omega_i \in \Omega_i$, the function f_{ω_i} is μ_j-integrable.
(2) Define $g_i : \Omega_i \to [-\infty, \infty]$ by

$$g_i(\omega_i) = \begin{cases} \int_{\Omega_j} f_{\omega_i} \, d\mu_j & \text{if the integral exists,} \\ 0 & \text{otherwise.} \end{cases}$$

The function g_i is μ_i-integrable and

$$\int_{\Omega_1 \times \Omega_2} f \, d\mu = \int_{\Omega_i} g_i \, d\mu_i.$$

The final equation is usually written as

$$\int_{\Omega_1 \times \Omega_2} f \, d\mu = \int_{\Omega_i} \int_{\Omega_j} f(\omega_1, \omega_2) \, d\mu_j \, d\mu_i.$$

Lemma 52. *If $R \in \mathcal{R}$ and $\mu(R) < \infty$, then $\chi_R \in (F)$.*

Lemma 53. *Suppose that (f_n) is a sequence of nonnegative μ-measurable functions that increases pointwise to a function f. If $f_n \in (F)$ for all n and $\sup_n \int f_n \, d\mu < \infty$, then $f \in (F)$.*

[Sketch. Monotone Convergence Theorem.]

Lemma 54. *If $f, g \in (F)$ and $c \in \mathbb{R}$, then $f + cg \in (F)$.*

Lemma 55. *Suppose that (f_n) is a sequence of nonnegative μ-measurable functions that decreases pointwise to a function f. If $f_n \in (F)$ for all n, then $f \in (F)$.*

[Sketch. Apply Lemma 53 to $f_1 - f_n$.]

Lemma 56. *Let $E \in \Sigma$ with $\mu(E) = 0$. Then $\chi_E \in (F)$.*

[Sketch. If $F = \cup R_n$, where $R_n \in \mathcal{R}$ and $\mu(F) < \infty$, then $\chi_F \in (F)$ by Lemmas 52 and 53. There is a sequence (F_n) so that $\mu(F_n) \to 0$, $F_n \supseteq F_{n+1} \supseteq E$, and each F_n is a countable union of sets in $\mathcal{R}$. By the above, $\chi_{F_n} \in (F)$. Let $F = \cap F_n$. Then $\chi_F \in (F)$ by Lemma 54. For μ_1 almost all ω_1, $(\chi_F)_{\omega_1}$ is μ_2 integrable with $\int_{\Omega_2} (\chi_F)_{\omega_1} \, d\mu_2 = 0$. Since $0 \leq (\chi_E)_{\omega_1} \leq (\chi_F)_{\omega_1}$ and μ_2 is complete, $(\chi_E)_{\omega_1}$ is μ_2-measurable and $\int_{\Omega_2} (\chi_E)_{\omega_1} \, d\mu_2 = 0$.]

Lemma 57. *If $E \in \Sigma$ and $\mu(E) < \infty$, then $\chi_E \in (F)$.*

[Sketch. E can be written as $F \setminus G$, where $G \subseteq F$, $\mu(G) = 0$ and $F = \cap F_n$, with $F_n \supseteq F_{n+1} \supseteq E$, and each F_n is a countable union of sets in $\mathcal{R}$.]

Theorem 58. *(Fubini's Theorem) Let $f : \Omega_1 \times \Omega_2 \to [-\infty, \infty]$ be a μ-integrable function. Then $f \in (F)$.*

[Sketch. Assume that $f \geq 0$. Choose a sequence of step functions $f_n \uparrow f$. Each f_n is μ-integrable. By Lemmas 56 and 57, $f_n \in (F)$ for all n. By Lemma 53, $f \in (F)$.]

Theorem 59. *(Tonelli's Theorem) Let $(\Omega_1, \Sigma_1, \mu_1)$ and $(\Omega_2, \Sigma_2, \mu_2)$ be complete σ-finite measure spaces and let $(\Omega_1 \times \Omega_2, \Sigma, \mu)$ be their product measure space. If f is a nonnegative μ-measurable function and $i, j \in \{1, 2\}$, $i \neq j$, then*

1. *For μ_i-almost all $\omega_i \in \Omega_i$, the function f_{ω_i} is μ_j-measurable.*
2. *Define $g_i : \Omega_i \to [-\infty, \infty]$ by*
$$g_i(\omega_i) = \begin{cases} \int_{\Omega_j} f_{\omega_i} \, d\mu_j & \text{if the integral exists,} \\ 0 & \text{otherwise.} \end{cases}$$
The function g_i is μ_i-measurable and
$$\int_{\Omega_1 \times \Omega_2} f \, d\mu = \int_{\Omega_i} g_i \, d\mu_i.$$

[Sketch. There are $R_1 \subseteq R_2 \subseteq \cdots$ in $\mathcal{R}$ so that $\Omega_1 \times \Omega_2 = \cup R_n$ and $\mu(R_n) < \infty$ for all n. Let $f_n = (f \wedge n)\chi_{R_n}$. Then f_n is μ-integrable. By Fubini, $f_n \in (F)$. In particular, (1) and (2) hold for f_n. Also $f_n \uparrow f$. Use the Monotone Convegence Theorem to prove conditions (1) and (2) for f.]

11. Fourier transforms

In this section, we consider the Lebesgue measure space $(\mathbb{R}^n, \Sigma, \lambda)$.

Proposition 60. *Suppose that $1 \le p \le \infty$. Let $f \in L^p(\mathbb{R}^n, \Sigma, \lambda)$ and $g \in L^1(\mathbb{R}^n, \Sigma, \lambda)$. For almost all $x \in \mathbb{R}^n$, the function $h(y) = f(x-y)g(y)$ is a Lebesgue integrable function of y on $\mathbb{R}^n$. Define*

$$(f * g)(x) = \begin{cases} \int f(x-y)g(y)\, d\lambda(y) & \text{if the integral exists,} \\ 0 & \text{otherwise.} \end{cases}$$

Then $f * g \in L^p(\mathbb{R}^n, \Sigma, \lambda)$ and $\|f * g\|_p \le \|f\|_p \|g\|_1$.

$f * g$ is called the **convolution** of f and g.

[Sketch. The case $p = \infty$ is easy. Consider the case $1 \le p < \infty$. Let $u(x)$ be any function in $L^q(\mathbb{R}^n, \Sigma, \lambda)$ with $\|u\|_q \le 1$. By Tonelli and Hölder,

$$\int \Big(\int |f(x-y)||g(y)|\, d\lambda(y) \Big) |u(x)|\, d\lambda(x) \le \|f\|_p \|u\|_q \int |g(y)|\, d\lambda \le \|f\|_p \|g\|_1.$$

In particular, $v(x) = \int |f(x-y)g(y)|\, dy < \infty$ for almost all x. Thus h is Lebesgue integrable for almost all x. Note that $|f * g(x)| \le v(x)$ for almost all x. By the equality in Hölder,

$$\|f * g\|_p \le \|v\|_p \le \|f\|_p \|g\|_1.]$$

Proposition 61. *Let $1 \le p \le \infty$, $f \in L^p(\mathbb{R}^n, \Sigma, \lambda)$ and $g, h \in L^1(\mathbb{R}^n, \Sigma, \lambda)$. Then $f * g = g * f$ and $f * (g * h) = (f * g) * h$ a.e.*

[Sketch. For the first equation, use Proposition 32. For the second equation, use Tonelli and Fubini.]

Denote by $C^\infty(\mathbb{R}^n)$ the space of all infinitely differentiable real-valued functions on $\mathbb{R}^n$. A real-valued function f on $\mathbb{R}^n$ is said to have **compact support** if $\{|f| \ne 0\}$ is a bounded set in $\mathbb{R}^n$. Let $C_c^\infty(\mathbb{R}^n)$ be the space of all functions in $C^\infty(\mathbb{R}^n)$ with compact support.

Proposition 62. *Let $f \in L^p(\mathbb{R}^n, \Sigma, \lambda)$ and $g \in C_c^\infty(\mathbb{R}^n)$. Then $f * g \in C^\infty(\mathbb{R}^n)$. If f also has compact support, then $f * g$ has compact support.*

[Sketch. For the first part, see Example 41.]

Definition. An **approximate identity** is a family of nonnegative functions $(K_\varepsilon)_{\varepsilon > 0}$ in $L^1(\mathbb{R}^n, \Sigma, \lambda)$ such that

(1) $\int K_\varepsilon\, d\lambda = 1$ for any $\varepsilon > 0$.
(2) For any $r > 0$, $\lim_{\varepsilon \to 0+} \int_{\{\|x\| > r\}} K_\varepsilon(x)\, d\lambda(x) = 0$.

Lemma 63. *Let $1 \le p < \infty$ and let $f \in L^p(\mathbb{R}^n, \Sigma, \lambda)$. For any $y \in \mathbb{R}^n$, define $f_y(x) = f(x-y)$. Then $\lim_{y \to 0} \|f - f_y\|_p = 0$.*

[Sketch. First prove it for $f = \chi_H$, where H is a bounded half-open interval. For the general case, approximate $f \in L^p$ by a linear combination of characteristic functions of bounded half-open intervals.]

The following result justifies the name "approximate identity".

Theorem 64. *Let $(K_\varepsilon)_{\varepsilon>0}$ be an approximate identity and let $f \in L^p(\mathbb{R}^n, \Sigma, \lambda)$, where $1 \le p < \infty$. Then $\lim_{\varepsilon \to 0+} \|f * K_\varepsilon - f\|_p = 0$.*

[Sketch. Use Tonelli to show that

$$\|f * K_\varepsilon - f\|_p \le \int \|f_y - f\|_p K_\varepsilon(y)\, d\lambda(y).$$

Split the integral into one part on $\{y : \|y\| \le r\}$ and another part on the rest of $\mathbb{R}^n$.]

Lemma 65. *Let K be a nonnegative Lebesgue integrable function such that $\int K\, d\lambda = 1$. Define $K_\varepsilon(x) = \frac{1}{\varepsilon^n} K(\frac{x}{\varepsilon})$ for any $\varepsilon > 0$. Then $(K_\varepsilon)_{\varepsilon>0}$ is an approximate identity.*

Proposition 66. *(Density of $C_c^\infty(\mathbb{R}^n)$ in $L^p(\mathbb{R}^n, \Sigma, \lambda)$) Let $1 \le p < \infty$ and let $f \in L^p(\mathbb{R}^n, \Sigma, \lambda)$. For any $\varepsilon > 0$, there exists $g \in C_c^\infty(\mathbb{R}^n)$ so that $\|f - g\|_p < \varepsilon$.*

[Sketch. Choose $0 \le K \in C_c^\infty(\mathbb{R}^n)$ so that $\int K\, d\lambda = 1$. Let $K_\varepsilon(x) = \frac{1}{\varepsilon^n} K(\frac{x}{\varepsilon})$. Then $K_\varepsilon \in C_c^\infty(\mathbb{R}^n)$. By Proposition 62, $f * K_\varepsilon \in C_c^\infty(\mathbb{R}^n)$. By Theorem 64, $\|f * K_\varepsilon - f\|_p \to 0$.]

If $x, y \in \mathbb{R}^n$, let $\langle x, y\rangle$ be the standard inner product of x and y.

Definition. Let $f \in L^1(\mathbb{R}^n, \Sigma, \lambda)$. Define $\widehat{f}$ on $\mathbb{R}^n$ by

$$\begin{aligned}\widehat{f}(y) &= \int_{\mathbb{R}^n} f(x) e^{-i\langle x,y\rangle}\, d\lambda(x)\\ &\overset{\text{def}}{=} \int_{\mathbb{R}^n} f(x)\cos\langle x,y\rangle\, d\lambda(x) - i\int_{\mathbb{R}^n} f(x)\sin\langle x,y\rangle\, d\lambda(x).\end{aligned}$$

$\widehat{f}$ is called the **Fourier transform** of f.

Example 67. *(Gaussian function) Let* $g(x) = \exp(-\frac{\|x\|^2}{2})$. *Then* $\widehat{g}(y) = (2\pi)^{\frac{n}{2}}\exp(-\frac{\|y\|^2}{2})$.

[Sketch. By Example 41 and integration by parts,

$$\frac{\partial\widehat{g}}{\partial y_k}(t_1, \ldots, t_n) = t_k \widehat{g}(t_1, \ldots, t_n).$$

Solve the differential equation with initial value $\widehat{g}(0) = (2\pi)^{\frac{n}{2}}$.]

Proposition 68. *Let* $f, g \in L^1(\mathbb{R}^n, \Sigma, \lambda)$. *Then*

(1) $\widehat{f}$ *is a bounded continuous function on* $\mathbb{R}^n$ *such that* $\sup_y |\widehat{f}(y)| \le \|f\|_1$.
(2) *(Riemann-Lebesgue Lemma)* $\lim_{\|y\|\to\infty} \widehat{f}(y) = 0$.
(3) $\widehat{f * g} = \widehat{f}\cdot\widehat{g}$ *a.e.*
(4) $\int_{\mathbb{R}^n} \widehat{f}\cdot g\, d\lambda = \int_{\mathbb{R}^n} f\cdot\widehat{g}\, d\lambda$.

[Sketch. (1) Continuity of $\widehat{f}$ follows from LDCT.
(2) In one dimension, first show that the result holds if $f \in C_c^\infty(\mathbb{R})$. Then use Proposition 66 and Proposition 68(1). In n dimensions, use the one dimensional result and Fubini.
(3) and (4) Use Tonelli and Fubini.]

The next two results show how to recover the function f from the Fourier transform $\widehat{f}$. The first formula is "natural" but the result only holds for certain $f \in L^1(\mathbb{R}^n, \Sigma, \lambda)$. The second result shows that it is always possible to recover f from $\widehat{f}$ by some form of averaging.

Theorem 69. *(Fourier Inversion Theorem) Let f be a Lebesgue integrable function on $\mathbb{R}^n$. If $\widehat{f}$ is Lebesgue integrable on $\mathbb{R}^n$, then*

$$f(x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \widehat{f}(y) e^{i\langle x,y\rangle}\, d\lambda(y) \text{ for almost all } x \in \mathbb{R}^n.$$

[Sketch. Define $g_\varepsilon(x) = \frac{1}{\varepsilon^n} g(\varepsilon x)$, where g is the Gaussian function. By Theorem 64, $\|f * g_\varepsilon - f\|_1 \to 0$. Use Example 67 to express g_ε in terms of $\widehat{g}$. Then apply Proposition 68(4) to compute $f * g_\varepsilon$. Deduce that

$$(f * g_\varepsilon)(x) \to \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \widehat{f}(y) e^{i\langle x,y\rangle}\, d\lambda(y).$$

Finally, by Proposition 43(2), there is a sequence $\varepsilon_k \downarrow 0$ so that $f * g_{\varepsilon_k} \to f$ a.e.]

Theorem 70. *(Fourier Summability Theorem) Let $(K_\varepsilon)_{\varepsilon>0}$ be an approximate identity. Assume that $\widehat{K_\varepsilon} \in L^1(\mathbb{R}^n)$ for all $\varepsilon > 0$. If $f \in L^1(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$, then the functions*

$$h_\varepsilon(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \widehat{f}(y)\widehat{K_\varepsilon}(y) e^{i\langle x,y\rangle}\, d\lambda(y)$$

converges in p-norm to f, i.e., $\lim_{\varepsilon\to 0+} \|h_\varepsilon - f\|_p = 0$.

[Sketch. Apply Fourier Inversion to $F * K_\varepsilon$.]

Proposition 71. *Let $\mathcal{F} : \mathbb{R}^n \to \mathbb{R}$ be given by*

$$\mathcal{F}(x) = \mathcal{F}(x_1, \ldots, x_n) = \frac{1}{(2\pi)^n} \prod_{k=1}^n \frac{\sin^2(x_k/2)}{(x_k/2)^2}.$$

Then $\mathcal{F} \in L^1(\mathbb{R}^n, \Sigma, \lambda)$ and

$$\widehat{\mathcal{F}}(y) = \widehat{\mathcal{F}}(y_1, \ldots, y_n) = \prod_{k=1}^n \max\{1 - |y_k|, 0\}.$$

Define $\mathcal{F}_\varepsilon(x) = \frac{1}{\varepsilon^n}\mathcal{F}(\frac{x}{\varepsilon})$. Then $(\mathcal{F}_\varepsilon)$ is an approximate identity. It is called **Fejér's kernel**.

[Sketch. Let $H(y) = \prod_{k=1}^n \max\{1 - |y_k|, 0\}$. By direct computation, $\widehat{H} = (2\pi)^n\mathcal{F}$. Use Fourier Inversion to show that $H = \widehat{\mathcal{F}}$.]

Corollary 72. *(Fejér's Theorem) Let $f \in L^1(\mathbb{R}^n)$. Define*

$$h_j(x) = \frac{1}{(2\pi)^n} \int_{[-j,j]^n} \widehat{f}(y) \prod_{k=1}^n (1 - \frac{|y_k|}{j}) e^{i\langle x,y\rangle}\, d\lambda(y).$$

Then $\lim_j \|h_j - f\|_1 = 0$.

12. Differentiation of monotone functions

An interval in $\mathbb{R}$ is said to be **nondegenerate** if it contains at least 2 points.

Definition. Let A be a set in $\mathbb{R}$. A collection $\mathcal{I}$ of nondegenerate closed bounded intervals in $\mathbb{R}$ is said to **cover** A **in the sense of Vitali** if for any $a \in A$ and any $\varepsilon > 0$, there exists $I \in \mathcal{I}$ such that $a \in I$ and $\lambda(I) < \varepsilon$.

Theorem 73. *(Vitali's Theorem) Let A be a set (not assumed measurable!) in $\mathbb{R}$ that is covered by a collection $\mathcal{I}$ of nondegenerate closed intervals in the sense of Vitali. Then there is a finite or countably infinite sequence (I_k) of pairwise disjoint members of $\mathcal{I}$ so that $\lambda(A \setminus (\cup I_k)) = 0$.*

[The proof uses a sort of "greedy algorithm" to select the intervals I_k one by one.]

Theorem 74. *Let I be an interval in $\mathbb{R}$ and let $f : I \to \mathbb{R}$ be a monotone function. Then f is differentiable almost everywhere on I.*

[Sketch. We may assume that I is a bounded open interval and f is nondecreasing. Fix $s < t$. Let $A_{s,t}$ be the set of points $x \in (a, b)$ so that

$$\liminf_{y\to x} \frac{f(y) - f(x)}{y - x} < s \text{ and } \limsup_{y\to x} \frac{f(y) - f(x)}{y - x} > t.$$

Use Theorem 73 to cover $A_{s,t}$ by a sequence of disjoint closed intervals $[u, v]$ where $|f(u) - f(v)| < s|u - v|$. Deduce that the rise of f over these intervals is at most approximately $s\lambda^*(A_{s,t})$. Repeat the argument using the "t" condition and deduce that the rise of f over these intervals is at least $t\lambda^*(A_{s,t})$. Contradiction. Then $\cup_{s,t\in\mathbb{Q};s<t} A_{s,t}$ has outer measure 0. f is differentiable outside of this set.]

It follows trivially from Theorem 74 that if $f : I \to \mathbb{R}$ is a difference of monotone functions, then f is differentiable a.e. on I. We now give an intrinsic characterization of such functions.

Definition. Let f be a real valued function defined on the closed bounded interval $[a, b]$. The **total variation** of f on $[a, b]$ is the quantity

$$\mathrm{Var}_{[a,b]} f = \sup\{\sum_{j=1}^{k} |f(x_j) - f(x_{j-1})| : k \in \mathbb{N}, a = x_0 < x_1 < \cdots < x_k = b\}.$$

f is said to be **of bounded variation** on $[a, b]$ if $\mathrm{Var}_{[a,b]} f < \infty$.

Example 75. *Let f be a Lebesgue integrable function on $[a, b]$. Define $F : [a, b] \to R$ by $F(x) = \int_a^x f \, d\lambda$. Then F has bounded variation on $[a, b]$.*

Proposition 76. *If f and g are of bounded variation on $[a, b]$ and $c \in \mathbb{R}$, then $f + g$, cf and fg are of bounded variation on $[a, b]$.*

Theorem 77. *Let $f : [a,b] \to \mathbb{R}$ be a function. Then f is of bounded variation on $[a,b]$ if and only if there are nondecreasing functions $f_1, f_2 : [a,b] \to \mathbb{R}$ so that $f = f_1 - f_2$. Moreover, f_1 and f_2 may be chosen so that $\operatorname{Var}_{[a,b]} f = f_1(b) - f_1(a) + f_2(b) - f_2(a)$.*

[Sketch. Let $V(x) = \operatorname{Var}_{[a,x]} f$ and set $f_1 = \frac{1}{2}(V + f)$, $f_2 = \frac{1}{2}(V - f)$.]

Theorem 78. *Let $f : [a,b] \to \mathbb{R}$ be a function of bounded variation on $[a,b]$. Then f is differentiable almost everywhere on $[a,b]$, f' is Lebesgue integrable on $[a,b]$ (f' is defined to be 0 at points where f is not differentiable), and $\int_a^b |f'|\, d\lambda \leq \operatorname{Var}_{[a,b]} f$.*

[Sketch. If $P = \{a = x_0 < \cdots < x_n = b\}$ is a parition of $[a,b]$, consider the derivative of the piecewise linear function that interpolates the points $(x_k, f(x_k))$. Take a sequence of partitions P_j so that $|P_j| \to 0$.]

Example 79. *Let f be the Cantor-Lebesgue function on $[0,1]$ (cf. Example 36). Then f is nondecreasing and hence f' exists a.e. in $[0,1]$. In fact, $f'(x) = 0$ a.e. Thus $\int_0^1 |f'|\, d\lambda = 0$. However, $\operatorname{Var}_{[0,1]} f = f(1) - f(0) = 1$.*

13. Lebesgue differentiation theorem

In this section, we are concerned with the Fundamental Theorem of Calculus for the Lebesgue integral.

Lemma 80. *Suppose that $f : [a,b] \to \mathbb{R}$ is Lebesgue integrable on the interval $[a,b]$. If $\int_a^x f\,d\lambda = 0$ for all $x \in [a,b]$, then $f = 0$ almost everywhere on $[a,b]$.*

[Consider $\mathcal{S}$, the set of all measurable subsets E of $[a,b]$ so that $\int_E f\,d\lambda = 0$. Then $\mathcal{S}$ is a σ-algebra on $[a,b]$ that includes all open sets in $[a,b]$ and all λ-null sets in $[a,b]$. Hence $\mathcal{S}$ contains all Lebesgue measurable subsets of $[a,b]$. Thus $\{f > 0\}$ and $\{f < 0\}$ both belong to $\mathcal{S}$.]

Theorem 81. *(Fundamental Theorem of Calculus I = Lebesgue differentiation theorem) Suppose that $f : [a,b] \to \mathbb{R}$ is Lebesgue integrable on the interval $[a,b]$. Then f is Lebesgue integrable on $[a,x]$ for all $x \in [a,b]$. Define $F : [a,b] \to \mathbb{R}$ by $F(x) = \int_a^x f\,d\lambda$. Then F is differentiable almost everywhere on $[a,b]$ and $F'(x) = f(x)$ for almost all $x \in [a,b]$.*

[Sketch. First suppose that $0 \le f \le M$ for some $M \in \mathbb{R}$ and let $F(x) = \int_a^x f\,d\lambda$. Use Theorem 78 on F and $M - F$ to show that $\int_a^x F'\,d\lambda = \int_a^x f\,d\lambda$ for any $x \in [a,b]$ and apply Lemma 80. For any general nonnegative f, consider $\min\{f, m\}$. For general f, consider f^+ and f^- separately.]

Theorem 82. *Suppose that I is an interval in $\mathbb{R}$ and $f : I \to \mathbb{R}$ is Lebesgue integrable on I. For almost every $x \in I$,*

$$\lim_{h\to 0+} \frac{1}{2h} \int_{x-h}^{x+h} |f(y) - f(x)|\,d\lambda(y) = 0.$$

A point x at which the last equation holds is called a **Lebesgue point** of the function f.

[Sketch. For any $r \in \mathbb{Q}$, use Theorem 81 to obtain a null set N_r such that

$$\lim_{h\to 0+} \frac{1}{2h} \int_{x-h}^{x+h} |f(y) - r|\,d\lambda(y) = 0$$

for all $x \notin N_r$. If $x \notin \cup_{r\in\mathbb{Q}} N_r$, choose a sequence (r_k) in $\mathbb{Q}$ converging to $f(x)$ and show that the result holds.]

Definition. Let $f : [a,b] \to \mathbb{R}$ be a function, where $a, b \in \mathbb{R}$. We say that f is **absolutely continuous** if for any $\varepsilon > 0$, there exists $\delta > 0$ such that if $k \in \mathbb{N}$, $a \le a_1 \le b_1 \le a_2 \le b_2 \le \cdots \le a_k \le b_k \le b$, and $\sum_{j=1}^k |b_j - a_j| \le \delta$, then $\sum_{j=1}^k |f(b_j) - f(a_j)| \le \varepsilon$.

Example 83. *Let $f : [a,b] \to \mathbb{R}$ be Lebesgue integrable on $[a,b]$. Define $F(x) = \int_a^x f\,d\lambda$ for all $x \in [a,b]$. Then F is absolutely continuous on $[a,b]$.*

Proposition 84. *Let $[a,b]$ a closed bounded interval in $\mathbb{R}$.*

(1) *An absolutely continuous function on* $[a, b]$ *is uniformly continuous.*
(2) *If* $f : [a, b] \to \mathbb{R}$ *is absolutely continuous, then it is of bounded variation on* $[a, b]$. *Hence* f *is differentiable almost everywhere and* f' *is Lesbegue integrable on* $[a, b]$.
(3) *If* f *and* g *are absolutely continuous on* $[a, b]$ *and* $c \in \mathbb{R}$, *then* $f + g$, cf *and* fg *are absolutely continuous on* $[a, b]$.

A function $f : [a, b] \to \mathbb{R}$ is **singular** if $f'(x)$ exists and is equal to 0 for almost all $x \in [a, b]$.

Proposition 85. *If* $f : [a, b] \to \mathbb{R}$ *is both absolutely continuous and singular, then* f *is constant.*

[Sketch. Let N be the set where $f'(x) \neq 0$ does not exist. Fix $\varepsilon > 0$. Use Theorem 73 to cover N^c by a countable union of intervals $[u, v]$ so that $|f(v) - f(u)| < \varepsilon|v - u|$. Choose finitely many of these intervals that cover "nearly" all of N^c. Form a partition using the endpoints of these finitely many intervals and the points a, b. Using the choice of the the intervals and the absolute continuity of f, show that $|f(b) - f(a)|$ is small. Taking ε as small as we please, it follows that $f(b) = f(a)$. The same argument applies to any $c \in (a, b)$.]

Theorem 86. *(Fundamental Theorem of Calculus II) Let* $[a, b]$ *be a closed interval in* $\mathbb{R}$ *and let* F *be a real valued function on* $[a, b]$. *The following are equivalent:*

(1) F' *is Lebesgue integrable on* $[a, b]$ *and* $F(x) = F(a) + \int_a^x F' \, d\lambda$ *for all* $x \in [a, b]$,
(2) *There is a Lebesgue integrable real valued function* f *on* $[a, b]$ *such that* $F(x) = F(a) + \int_a^x f \, d\lambda$ *for all* $x \in [a, b]$,
(3) F *is absolutely continuous on* $[a, b]$.

[Sketch. (3) $\implies$ (1). Apply Proposition 85 to the function $G(x) = F(x) - \int_a^x F' \, d\lambda$.]

Theorem 87. *(Lebesgue decomposition) Let* $f : [a, b] \to \mathbb{R}$ *be a function of bounded variation. Then* $f = g + h$, *where* g *is absolutely continuous on* $[a, b]$ *and* h *is singular on* $[a, b]$. *The decomposition is unique up to additive constants.*

[Sketch. Take $g(x) = \int_a^x f' \, d\lambda$ and $h = f - g$.]