MATH454 - Analysis 3

Measure spaces; Integration.

Based on lectures from Fall 2024 by Prof. Linan Chen.
Notes by Louis Meunier

Contents

§1 Sigma Algebras and Measures

§1.1 A Review of Riemann Integration

Let $f : \mathbb{R} \to \mathbb{R}$ and $[a, b] \subset \mathbb{R}$. Define a **partition** of $[a, b]$ as the set

$$\operatorname{part}([a, b]) := \{a =: x_0 < x_1 < \cdots < x_N := b\}.$$

We can then define the upper and lower Riemann integrals of f over the region $[a, b]$ as

$$\text{upper:} \qquad \overline{\int_a^b} f(x)\,\mathrm{d}x := \inf_{\operatorname{part}([a,b])} \left\{ \sum_{i=1}^{N} \sup_{x \in [x_{i-1}, x_i]} f(x) \cdot (x_i - x_{i-1}) \right\}$$

$$\text{lower:} \qquad \underline{\int_a^b} f(x)\,\mathrm{d}x := \sup_{\operatorname{part}([a,b])} \left\{ \sum_{i=1}^{N} \inf_{x \in [x_{i-1}, x_i]} f(x) \cdot (x_i - x_{i-1}) \right\}.$$

We then say f **Riemann integrable** if these two quantities are equal, and denote this value by $\int_a^b f(x)\,\mathrm{d}x$.

Many "nice-enough" (continuous, monotonic, etc.) functions are Riemann integrable, but many that we would like to be able to "integrate" are simply not, for instance Dirichlet's function $x \mapsto \begin{cases} 1 x \in \mathbb{Q} \setminus [a,b] \\ 0 x \in \mathbb{Q}^c \setminus [a,b] \end{cases}$. Hence, we need a more general notion of integration.

§1.2 Sigma Algebras

↪**Definition 1.1** (Sigma algebra): Let X be a *space* (a nonempty set) and $\mathcal{F}$ a collection of subsets of X. $\mathcal{F}$ a *sigma algebra* or simply σ-algebra of X if the following hold:

1. $X \in \mathcal{F}$
2. $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$ (closed under complement)
3. $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ (closed under countable unions)

↪**Proposition 1.1**:

4. $\varnothing \in \mathcal{F}$
5. $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{F} \Rightarrow \bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$
6. $A_1, \ldots, A_n \in \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n, \bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$
7. $A, B \in \mathcal{F} \Rightarrow A \setminus B, B \setminus A \in \mathcal{F}$

⊛ **Example 1.1**: The "largest" sigma algebra of a set X is the power set 2^X, the smallest the trivial $\{\varnothing, X\}$.

Given a set $A \subset X$, the set $\mathcal{F}_A := \{\varnothing, X, A, A^c\}$ is a sigma algebra; given two disjoint sets $A, B \subset X$, then $\mathcal{F}_{A,B} := \{\varnothing, X, A, A^c, B, B^c, A \cup B, A^c \cap B^c\}$ a sigma algebra.

↪**Definition 1.2** (Generating a sigma algebra): Let X be a nonempty set, and $\mathcal{C}$ a collection of subsets of X. Then, the σ-algebra *generated* by $\mathcal{C}$, denoted $\sigma(\mathcal{C})$, is such that

1. $\sigma(\mathcal{C})$ a sigma algebra with $\mathcal{C} \subseteq \sigma(\mathcal{C})$
2. if $\mathcal{F}'$ a sigma algebra with $\mathcal{C} \subseteq \mathcal{F}'$, then $\mathcal{F}' \supseteq \sigma(\mathcal{C})$

Namely, $\sigma(\mathcal{C})$ is the smallest sigma algebra "containing" (as a subset) $\mathcal{C}$.

↪**Proposition 1.2**:

1. $\sigma(\mathcal{C}) = \bigcap\{\mathcal{F} : \mathcal{F} \text{ a sigma algebra containing } \mathcal{C}\}$
2. if $\mathcal{C}$ itself a sigma algebra, then $\sigma(\mathcal{C}) = \mathcal{C}$
3. if $\mathcal{C}_1, \mathcal{C}_2$ are two collections of subsets of X such that $\mathcal{C}_1 \subseteq \mathcal{C}_2$, then $\sigma(\mathcal{C}_1) \subseteq \sigma(\mathcal{C}_2)$

↪**Definition 1.3** (The Borel sigma-algebra): The *Borel σ-algebra*, denoted $\mathfrak{B}_{\mathbb{R}}$, on the real line is given by

$$\mathfrak{B}_{\mathbb{R}} := \sigma(\{\text{open subsets of } \mathbb{R}\}).$$

We call sets in $\mathfrak{B}_{\mathbb{R}}$ *Borel sets*.

↪**Proposition 1.3**: $\mathfrak{B}_{\mathbb{R}}$ is also generated by the sets

- $\{(a,b) : a < b \in \mathbb{R}\}$
- $\{(a,b] : a < b \in \mathbb{R}\}$
- $\{[a,b] : a < b \in \mathbb{R}\}$
- $\{[a,b) : a < b \in \mathbb{R}\}$ $\circledast$
- $\{(-\infty,c) : c \in \mathbb{R}\}$
- $\{(-\infty,c] : c \in \mathbb{R}\}$
- etc.

Proof. We prove just $\circledast$. It suffices to show that the generating sets of each σ-algebra is contained in the other σ-algebra. Let $a < b \in \mathbb{R}$. Then,

$$(a,b) = \bigcup_{n=1}^{\infty} \underbrace{\left[a + \frac{1}{n}, b\right)}_{\in \circledast} \in \sigma(\{[a,b)\}) \Rightarrow \mathfrak{B}_{\mathbb{R}} \subseteq \sigma(\{[a,b)\}).$$

Conversely,

$$[a,b) = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, b\right) \in \mathfrak{B}_{\mathbb{R}}.$$

■

↪**Proposition 1.4**: All intervals (open, closed, half open, half closed, finite, etc) are Borel sets; any set obtained from countable set operations of intervals are Borel; all singletons are Borel; any finite and countable sets are Borel.

§1.3 Measures

↪**Definition 1.4** (Measurable Space): Let X be a space and $\mathcal{F}$ a σ-algebra. We call the tuple $(X, \mathcal{F})$ a *measurable space*.

↪**Definition 1.5** (Measure): Let $(X, \mathcal{F})$ be a measurable space. A *measure* is a function $\mu : \mathcal{F} \to [0, \infty]$ satisfying

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

(ii) if $\{A_n\} \subseteq \mathcal{F}$ a sequence of (pairwise) disjoint sets, then

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n),$$

i.e. μ is *countably additive*. We further call μ

- *finite* if $\mu(X) < \infty$,
- a *probability measure* if $\mu(X) = 1$,
- σ-finite if $\exists \{A_n\} \subseteq \mathcal{F}$ such that $X = \bigcup_{n=1}^{\infty} A_n$ with $\mu(A_n) < \infty \; \forall\, n \geq 1$,

and call the triple $(X, \mathcal{F}, \mu)$ a *measure space*.

⊛ **Example 1.2**: The measure on $\mathfrak{B}_{\mathbb{R}}$ given by

$$A \mapsto \begin{cases} |A| \text{ if } A \text{ finite} \\ \infty \quad \text{else} \end{cases}$$

is called the *counting measure*.

Fix $x_0 \in \mathbb{R}$, then the measure on $\mathfrak{B}_{\mathbb{R}}$ given by

$$A \mapsto \begin{cases} 1 \text{ if } x_0 \in A \\ 0 \text{ else} \end{cases}$$

is called the *point mass at x_0*.

↪**Theorem 1.1** (Properties of Measures): Fix a measure space $(X, \mathcal{F}, \mu)$. The following properties hold:

1. (finite additivity) For any sequence $\{A_n\}_{n=1}^N \subseteq \mathcal{F}$ of disjoint sets,
$$\mu\left(\bigcup_{n=1}^N A_n\right) = \sum_{n=1}^N \mu(A_n).$$
2. (monotonicity) For any $A \subseteq B \in \mathcal{F}$, then $\mu(A) \leq \mu(B)$.
3. (countable/finite subadditivity) For any sequence $\{A_n\} \subseteq \mathcal{F}$ (**not** necessarily disjoint),
$$\mu\left(\bigcup_{n=1}^\infty A_n\right) \leq \sum_{n=1}^\infty \mu(A_n),$$
an analogous statement holding for a finite collection of sets $A_1, ..., A_N$.
4. (continuity from below) For $\{A_n\} \subseteq \mathcal{F}$ such that $A_n \subseteq A_{n+1}\ \forall\, n \geq 1$ (in which case we say $\{A_n\}$ "increasing" and write $A_n \uparrow$) we have
$$\mu\left(\bigcup_{n=1}^\infty A_n\right) = \lim_{n \to \infty} \mu(A_n).$$
5. (continuity from above) For $\{A_n\} \subseteq \mathcal{F}, A_n \supseteq A_{n+1}\ \forall\, n \geq 1$ (we write $A_n \downarrow$) we have that **if** $\mu(A_1) < \infty$,
$$\mu\left(\bigcap_{n=1}^\infty A_n\right) = \lim_{n \to \infty} \mu(A_n).$$

Remark 1.1: In 4., note that since A_n increasing, that the union $\bigcup_{n=1}^\infty A_n \supseteq A_m$ for any arbitrarily large m; indeed, one could logically right $\lim_{n\to\infty} A_n = \bigcup_{n=1}^\infty A_n$. In this notation, then, 4. simply states that we may interchange limit and measure. A similar argument can be viewed for 5. (how?).

Remark 1.2: The finiteness condition in 5. may be slightly modified such as to state that $\mu(A_n) < \infty$ for some n; remark why this would suffice to ensure the entire rest of the sequence has finite measure.

PROOF.

1. Extend $A_1, ..., A_N$ to an infinite sequence by $A_n := \emptyset$ for $n > N$. Then this simply follows from countable additivity and $\mu(\emptyset) = 0$.
2. We may write $B = A \cup (B \setminus A)$; this is a disjoint union of sets. By finite additivity, then,
$$\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A),$$
since the measure is positive.
3. We prove only for a countable union; use the technique from 1. to extend to finite. We first "disjointify" the sequence such that we can use the countable additivity

axiom. Let $B_1 = A_1, B_n = A_n \setminus \left(\bigcup_{i=1}^{n-1} A_i\right)$ for $n \geq 2$. Remark then that $\{B_n\} \subseteq \mathcal{F}$ is a disjoint sequence of sets, and that $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$. By countable additivity and subadditivity,

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mu(B_n) \leq \sum_{n=1}^{\infty} \mu(A_n).$$

4. We again "disjointify" the sequence $\{A_n\}$. Put $B_1 = A_1$, $B_n = A_n \setminus A_{n-1}$ for all $n \geq 2$ (remark that this is equivalent to the construction from the previous proof because the sets are increasing). Then, again, $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$, and in particular, for all $N \geq 1$, $\bigcup_{n=1}^{N} B_n = A_N$. Then

$$\begin{aligned}\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=1}^{\infty} A_n\right) &= \sum_{n=1}^{\infty} \mu(B_n) \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} \mu(B_n) \\ &= \lim_{N\to\infty} \mu\left(\bigcup_{n=1}^{N} B_n\right) = \lim_{N\to\infty} \mu(A_N).\end{aligned}$$

5. We yet again disjointify, backwards (in a way) from the previous case. Put $B_n = A_1 \setminus A_n$ for all $n \geq 1$. Then, $\{B_n\} \subseteq \mathcal{F}$, B_n increasing, and $\bigcup_{n=1}^{\infty} B_n = A_1 \setminus \bigcap_{n=1}^{\infty} A_n$. Then, by continuity from below,

$$\mu\left(A_1 \setminus \bigcap_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} \mu(B_n) = \lim_{n\to\infty} \mu(A_1 \setminus A_n)$$

and also

$$\begin{aligned}\mu(A_1) &= \mu\left(A_1 \setminus \bigcap_{n=1}^{\infty} A_n\right) + \mu\left(\bigcap_{n=1}^{\infty} A_n\right) \\ &= \mu(A_1 \setminus A_n) + \mu(A_n),\end{aligned}$$

and combining these two equalities yields the desired result.

■

§1.4 Constructing the Lebesgue Measure on $\mathbb{R}$

↪**Definition 1.6** (Lebesgue outer measure): For all $A \subseteq \mathbb{R}$, define

$$m^*(A) := \inf\left\{\sum_{n=1}^{\infty} \ell(I_n) : A \subseteq \bigcup_{n=1}^{\infty} I_n, I_n \text{ open intervals}\right\},$$

called the *Lebesgue outer measure* of A (where $\ell(I)$ is the length of interval I, i.e. the absolute value of the difference of its endpoints, if finite, or ∞ if not).

↪**Proposition 1.5**: The following properties of m^* hold:

1. $m^*(A) \geq 0$ for all $A \subseteq \mathbb{R}$, and $m^*(\varnothing) = 0$.
2. (monotonicity) For $A \subseteq B$, $m^*(A) \leq m^*(B)$.
3. (countable subadditivity) For $\{A_n\}, A_n \subseteq \mathbb{R}$, $m^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} m^*(A_n)$.[1]
4. If $I \subseteq \mathbb{R}$ an interval, then $m^*(I) = \ell(I)$.
5. m^* is translation invariant; for any $A \subseteq R, x \in \mathbb{R}$, $m^*(A) = m^*(A+x)$ where $A + x := \{a + x : a \in A\}$.
6. For all $A \subseteq \mathbb{R}$, $m^*(A) = \inf\{m^*(B) : A \subseteq B \subseteq R, B - \text{open}\}$.
7. If $A = A_1 \cup A_2 \subseteq \mathbb{R}$ with $d(A_1, A_2) > 0$,[2] then $m^*(A_1) + m^*(A_2) = m^*(A)$.
8. If $A = \bigcup_{k=1}^{\infty} J_k$ where J_k's are "almost disjoint intervals" (i.e. share at most endpoints), then $m^*(A) = \sum_{k=1}^{\infty} m^*(J_k) = \sum_{k=1}^{\infty} \ell(J_k)$.

Proof.

3. If $m^*(A_n) = \infty$, for any n, we are done, so assume wlog $m^*(A_n) < \infty$ for all n. Then, for each n and $\varepsilon > 0$, one can choose open intervals $\{I_{n,i}\}_{i \geq 1}$ such that $A_n \subseteq \bigcup_{i=1}^{\infty} I_{n,i}$ and $\sum_{i=1}^{\infty} \ell(I_{n,i}) \leq m^*(A_n) + \frac{\varepsilon}{2^n}$. Hence
$$\bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1, i=1}^{\infty} I_{n,i}$$
$$\Rightarrow m^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n,i=1}^{\infty} \ell(I_{n,i}) = \sum_{n=1}^{\infty}\sum_{i=1}^{\infty} \ell(I_{n,i}) \leq \sum_{n=1}^{\infty}\left(m^*(A_n) + \frac{\varepsilon}{2^n}\right) = \sum_{n=1}^{\infty} m^*(A_n) + \varepsilon,$$

and as ε arbitrary, the statement follows.

4. We prove first for $I = [a, b]$. For any $\varepsilon > 0$, set $I_1 = (a - \varepsilon, b + \varepsilon)$; then $I \subseteq I_1$ so $m^*(I) \leq \ell(I_1) = (b-1) + 2\varepsilon$ hence $m^*(I) \leq b - a = \ell(I)$. Conversely, let $\{I_n\}$ be any open-interval convering of I (wlog, each of finite length; else the statement holds trivially). Since I compact, it can be covered by finitely many of the I_n's, say $\{I_n\}_{n=1}^{N}$, denoting $I_n = (a_n, b_n)$ (with relabelling, etc). Moreover, we can pick the a_n, b_n's such that $a_1 < a, b_N > b$, and generally $a_n < b_{n-1}\ \forall\, 2 \leq n \leq N$. Then,
$$\begin{aligned}\sum_{n=1}^{\infty} \ell(I_n) \geq \sum_{n=1}^{N} \ell(I_n) &= b_1 - a_1 + \sum_{n=2}^{N}(b_n - a_n) \\ &\geq b_1 - a_1 + \sum_{n=2}^{N}(b_n - b_{n-1}) \\ &= b_N - a_1 \geq b - 1 = \ell(I),\end{aligned}$$
hence since the cover was arbitrary, $m^*(A) \geq \ell(I)$, and equality holds.

Now, suppose I finite, with endpoints $a < b$. Then for any $\frac{b-a}{2} > \varepsilon > 0$, then
$$[a + \varepsilon, b - \varepsilon] \subseteq I \subseteq [a - \varepsilon, b + \varepsilon],$$

[1]More generally, any set function on $2^{\mathbb{R}}$ that satisfies 1., 2., and 3. is called an *outer measure*.
[2]Remark: this is a stronger requirement than disjointness!

hence by monotonicity and the previous part of this proof

$$m^*([a+\varepsilon, b-\varepsilon]) = b-a-2\varepsilon \leq m^*(I) \leq b-a+2\varepsilon = m^*([a-\varepsilon, b+\varepsilon]),$$

from which it follows that $m^*(I) = b-a = \ell(I)$.

Finally, suppose I infinite. Then, $\forall\, M \geq 0, \exists$ closed, finite interval I_M with $I_M \subseteq I$ and $\ell(I_M) \geq M$. Hence, $m^*(I) \geq m^*(I_M) \geq M$ and thus as M arbitrary it must be that $m^*(I) = \infty = \ell(I)$.

6. Denote $\tilde{m}(A) := \inf\{m^*(B) : A \subseteq B \subseteq \mathbb{R}, B - \text{open}\}$. For any $A \subseteq B \subseteq \mathbb{R}$ with B open, monotonicity gives that $m^*(A) \leq m^*(B)$, hence $m^*(A) \leq \tilde{m}(A)$. Conversely, assuming wlog $m^*(A) < \infty$ (else holds trivially), then for all $\varepsilon > 0$, there exists $\{I_n\}$ such that $A \subseteq \bigcup_{n=1}^\infty I_n$ with $\sum_{n=1}^\infty \ell(I_n) \leq m^*(A) + \varepsilon$. Setting $B := \bigcup_{n=1}^\infty I_n$, we have that $A \subseteq B$ and $m^*(B) = m^*(\bigcup I_n) \leq$ (by finite subadditivity) $\sum_{n=1}^\infty m^*(I_n) = \sum_{n=1}^\infty \ell(I_n) \leq m^*(A) + \varepsilon$ hence $m^*(B) \leq m^*(A)$ for all B. Thus $m^*(A) \geq \tilde{m}(A)$ and equality holds.

7. Put $\delta := d(A_1, A_2) > 0$. Clearly $m^*(A) \leq m^*(A_1) + m^*(A_2)$ by finite subadditivity. wlog, $m^*(A) < \infty$ (and hence $m^*(A_i) < \infty, i = 1,2$) (else holds trivially). Then $\forall\, \varepsilon > 0, \exists\, \{I_n\} : A \subseteq \bigcup I_n$ and $\sum \ell(I_n) \leq m^*(A) + \varepsilon$. Then, for all n, we consider a "refinement" of I_n; namely, let $\{I_{n,i}\}_{i \geq 1}$ such that $I_n \subseteq \bigcup_i I_{n,i}$ and $\ell\big(I_{n,i}\big) < \delta$ and $\sum_i \ell\big(I_{n,i}\big) \leq \ell(I_n) + \frac{\varepsilon}{2^n}$. Relabel $\{I_{n,i} : n, i \geq 1\} \rightsquigarrow \{J_m : m \geq 1\}$ (both are countable). Then, $\{J_m\}$ defines an open-interval cover of A, and since $\ell(J_m) < \delta$ for each m, J_m intersects at most one A_i. For each m and $p = 1, 2$, put

$$M_p := \Big\{m : J_m \cap A_p \neq \varnothing\Big\},$$

noting that $M_1 \cap M_2 = \varnothing$. Then $\Big\{J_m : m \in M_p\Big\}$ is an open covering of A_p, and so

$$\begin{aligned} m^*(A_1) + m^*(A_2) &\leq \sum_{m \in M_1} \ell(J_m) + \sum_{m \in M_2} \ell(J_m) \\ &\leq \sum_{m=1}^\infty \ell(J_m) = \sum_{n,i=1}^\infty \ell(I_n, i) \\ &\leq \sum_n \Big(\ell(I_n) + \frac{\varepsilon}{2^n}\Big) \\ &= \sum_n \ell(I_n) + \varepsilon \\ &\leq m^*(A) + 2\varepsilon, \end{aligned}$$

and hence equality follows.

8. If $\ell(J_k) = \infty$ for some k, then since $J_k \subseteq A$, subadditivity gives us that $m^*(J_k) \leq m^*(A)$ and so $m^*(A) = \infty = \sum_{k=1}^\infty \ell(J_k)$ (since if any J_k infinite, the sum of the lengths of all of them will also be infinite).

Suppose then $\ell(J_k) < \infty$ for all k. Fix $\varepsilon > 0$. Then for all $k \geq 1$, choose $I_k \subseteq J_k$ such that $\ell(J_k) \leq \ell(I_k) + \frac{\varepsilon}{2^k}$. For any $N \geq 1$, we can choose a subset $\{I_1, ..., I_N\}$ of intervals such that all are disjoint, with strictly positive distance between them, and so

$$\bigcup_{k=1}^{N} I_k \subseteq \bigcup_{k=1}^{N} I_k \subseteq A$$

$$\begin{aligned}
\Rightarrow m^*(A) \geq m^*\left(\bigcup_{k=1}^{N} I_k\right) &\geq \sum_{k=1}^{N} \ell(I_k) \\
&\geq \sum_{k=1}^{N} \left(\ell(J_k) - \frac{\varepsilon}{2^k}\right) \\
&\geq \sum_{k=1}^{N} \ell(J_k) - \varepsilon
\end{aligned}$$

$$\Rightarrow m^*(A) \geq \sum_{k=1}^{\infty} \ell(J_k),$$

the second inequality following from finite subadditivity. The converse of the final inequality holds trivially. ■

§1.5 Lebesgue-Measurable Sets

↪**Definition 1.7**: $A \subseteq \mathbb{R}$ is *m^*-measurable* if $\forall B \subseteq \mathbb{R}$,

$$m^*(B) = m^*(B \cap A) + m^*(B \cap A^c).$$

Remark 1.3: By subadditivity, $\leq$ always holds in the definition above.

↪**Theorem 1.2** (Carathéodary's Theorem): Let

$$\mathcal{M} := \{A \subseteq \mathbb{R} : A \;\; m^* - \text{measurable}\}.$$

Then, $\mathcal{M}$ is a σ-algebra of subsets of $\mathbb{R}$.

Define $m : \mathcal{M} \to [0, \infty]$, $m(A) = m^*(A)$. Then, m is a measure on $\mathcal{M}$, called the *Lebesgue measure* on $\mathbb{R}$. We call sets in $\mathcal{M}$ *Lebesgue-measurable* or simply *measurable* (if clear from context) accordingly. We call $(\mathbb{R}, \mathcal{M}, m)$ the *Lebesgue measure space*.

Proof. The first two σ-algebra axioms are easy. We have for any $B \subseteq \mathbb{R}$ that

$$m^*(B \cap \mathbb{R}) + m^*(B \cap \mathbb{R}^c) = m^*(B) + m^*(B \cap \varnothing) = m^*(B)$$

so $\mathbb{R} \in \mathcal{M}$. Further, $A \in \mathcal{M} \Rightarrow A^c \in \mathcal{M}$ by the symmetry of the requirement for sets to be in $\mathcal{M}$.

The final axiom takes more work. We show first $\mathcal{M}$ closed under finite unions; by induction it suffices to show for 2 sets. Let $A_1, A_2 \in \mathcal{M}$. Then, for all $B \subseteq \mathbb{R}$,

$$\begin{aligned}
m^*(B) &= m^*(B \cap A_1) + m^*(B \cap A_1^c) \\
&= m^*(B \cap A_1) + m^*(B \cap A_1^c \cap A_2) + m^*(B \cap A_1^c \cap A_2^c) \\
&= m^*(B \cap A_1) + m^*(B \cap A_1^c \cap A_2) + m^*\left(B \cap (A_1 \cup A_2)^c\right)
\end{aligned}$$

Note that $(B \cap A_1) \cup (B \cap A_1^c \cap A_2^c) = B \cap (A_1 \cup A_2)$, hence by subadditivity,

$$m^*(B) \geq m^*(B \cap (A_1 \cup A_2)) + m^*\big(B \cap (A_1 \cup A_2)^c\big),$$

and since the other direction of the inequality comes for free, we conclude $A_1 \cup A_2 \in \mathcal{M}$.

Let now $\{A_n\} \subseteq \mathcal{M}$. We "disjointify" $\{A_n\}$; put $B_1 := A_1$, $B_n := A_n \setminus \bigcup_{i=1}^{n-1} A_i$, $n \geq 2$, noting $\bigcup_n A_n = \bigcup_n B_n$, and each $B_n \in \mathcal{M}$, as each is but a finite number of set operations applied to the A_n's, and thus in $\mathcal{M}$ as demonstrated above. Put $E_n := \bigcup_{i=1}^n B_i$, noting again $E_n \in \mathcal{M}$. Then, for all $B \subseteq \mathbb{R}$,

$$\begin{aligned}
m^*(B) &= m^*\left(\underbrace{B \cap E_n}_{\text{chop up } B_n}\right) + m^*\left(\underbrace{B \cap E_n^c}_{E_n \subseteq \cup B_n \Rightarrow E_n^c \supseteq (\cup B_n)^c}\right) \\
&\geq m^*\left(B \cap \underbrace{E_n \cap B_n}_{=B_n}\right) + m^*\left(B \cap \underbrace{E_n \cap B_n^c}_{=E_{n-1}}\right) + m^*\left(B \cap \left(\bigcup_{n=1}^\infty B_n\right)^c\right) \\
&\geq m^*(B \cap B_n) + m^*\left(\underbrace{B \cap E_{n-1}}_{\text{chop up } B_{n-1}}\right) + m^*\left(B \cap \left(\bigcup_{n=1}^\infty B_n\right)^c\right) \\
&\geq m^*(B \cap B_n) + m^*(B \cap E_{n-1} \cap B_{n-1}) \\
&\qquad + m^*(B \cap E_{n-1} \cap B_{n-1}^c) + m^*\left(B \cap \left(\bigcup_{n=1}^\infty B_n\right)^c\right).
\end{aligned}$$

Notice that the last line is essentially the second applied to B_{n-1}; hence, we have a repeating (essentially, "descending") pattern in this manner, which we repeat until $n \to 1$. We have, thus, that

$$m^*(B) \geq \sum_{i=1}^n [m^*(B \cap B_i)] + m^*\left(B \cap \left(\bigcup_{n=1}^\infty B_n\right)^c\right),$$

so taking $n \to \infty$,

$$\begin{aligned}
m^*(B) &\geq \sum_{i=1}^\infty [m^*(B \cap B_i)] + m^*\left(B \cap \left(\bigcup_{n=1}^\infty B_n\right)^c\right) \\
&\geq m^*\left(B \cap \left(\bigcup_{n=1}^\infty B_n\right)\right) + m^*\left(B \cap \left(\bigcup_{n=1}^\infty B_n\right)^c\right).
\end{aligned}$$

As usual, the inverse inequality comes for free, and thus we can conclude $\bigcup_{n=1}^\infty B_n$ also m^*-measurable, and thus so is $\bigcup_{n=1}^\infty A_n$. This proves $\mathcal{M}$ a σ-algebra.

We show now m a measure. By previous propositions, we have that $m \geq 0$ and $m(\varnothing) = 0$ (since $m = m^* \mid_{\mathcal{M}}$), so it remains to prove countable additivity.

Let $\{A_n\} \subseteq \mathcal{M}$-disjoint. Following precisely the same argument as above, used to prove that $\mathcal{M}$ closed under countable unions, shows that for any $n \geq 1$

$$m\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} m(A_i),$$

that is, finite additivity holds, and thus by subadditivity

$$m\left(\bigcup_{i=1}^{\infty} A_i\right) \geq m\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} m(A_i),$$

and so taking the limit of $n \to \infty$, we have

$$m\left(\bigcup_{i=1}^{\infty} A_i\right) \geq \sum_{i=1}^{\infty} m(A_i),$$

with the converse inequality coming for free. Thus, m indeed a measure on $\mathcal{M}$. ■

↪**Proposition 1.6**: $\mathcal{M}, m$ translation invariant; for all $A \in \mathcal{M}, x \in \mathbb{R}, x + A = \{x + a : a \in A\} \in \mathcal{M}$ and $m(A) = m(A + x)$.

Remark 1.4: We would like this to hold, heuristically, since if we shift sets on the real line, we should expect their length to remain constant.

Proof. For all $B \subseteq \mathbb{R}$, we have (since m^* translation invariant)

$$\begin{aligned} m^*(B) = m^*(B - x) &= m^*\left(\underbrace{(B - x) \cap A}_{=B \cap (A+x)}\right) + m^*\left(\underbrace{(B - x) \cap A^c}_{=B \cap (A^c + x) = B \cap (A+x)^c}\right) \\ &= m^*(B \cap (A + x)) + m^*\left(B \cap (A + x)^c\right), \end{aligned}$$

thus $A + x \in \mathcal{M}$, and since m^* translation invariant, it follows that m is. ■

↪**Theorem 1.3**: $\forall\, a, b \in \mathbb{R}$ with $a < b$, $(a, b) \in \mathcal{M}$, and $m((a, b)) = b - a$.

Remark 1.5: Again, we'd like this to hold, heuristically, since we would like the measure of an interval to simply be its length; we'd moreover like to be able to measure intervals, i.e. have intervals be contained in $\mathcal{M}$.

↪**Corollary 1.1**: $\mathfrak{B}_{\mathbb{R}} \subseteq \mathcal{M}$

Proof. $\mathfrak{B}_{\mathbb{R}}$ is generated by open intervals of the form (a, b). All such intervals are in $\mathcal{M}$ by the previous theorem, and hence the proof. ■

§1.6 Properties of the Lebesgue Measure

↪**Proposition 1.7** (Regularity Properties of m): For all $A \in \mathcal{M}$, the following hold.
- For all $\varepsilon > 0, \exists\, G$ open such that $A \subseteq G$ and $m(G \setminus A) < \varepsilon$.
- For all $\varepsilon > 0$, $\exists\, F$-closed such that $F \subseteq A$ and $m(A \setminus F) \leq \varepsilon$.
- $m(A) = \inf\{m(G) : G \text{ open}, G \supseteq A\}$.
- $m(A) = \sup\{m(K) : K \text{ compact}, K \subseteq A\}$.
- If $m(A) < \infty$, then for all $\varepsilon > 0$, $\exists\, K \subseteq A$ compact, such that $m(A \setminus K) < \varepsilon$.
- If $m(A) < \infty$, then for all $\varepsilon \geq 0$, $\exists$ finite collection of open intervals $I_1, ..., I_N$ such that $m\left(A \triangle \left(\bigcup_{n=1}^N I_n\right)\right) \leq \varepsilon$.

↪**Proposition 1.8** (Completeness of m): $(\mathbb{R}, \mathcal{M}, m)$ is *complete*, in the sense that for all $A \subseteq \mathbb{R}$, if $\exists\, B \in \mathcal{M}$ such that $A \subseteq B$ and $m(B) = 0$, then $A \in \mathcal{M}$ and $m(A) = 0$.

Equivalently, any subset of a null set is again a null set.

Remark 1.6: In general, $A \in \mathcal{F}, B \subseteq A \not\Rightarrow B \in \mathcal{F}$.

↪**Proposition 1.9**: Up to rescaling, m is the unique, nontrivial measure on $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$ that is finite on compact sets and is translation invariant, i.e. if μ another such measure on $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$ with $\mu = c \cdot m$ for $c > 0$, then $\mu = m$.

Remark 1.7: Such a c is simply $c = \mu((0, 1))$.

To prove this proposition, we first introduce some helpful tooling:

↪**Theorem 1.4** (Dynkin's π-d): Given a space X, let $\mathcal{C}$ be a collection of subsets of X. $\mathcal{C}$ is called a *π-system* if $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$ (that is, it is closed under finite intersections).

Let $\mathcal{F} = \sigma(\mathcal{C})$, and suppose μ_1, μ_2 are two finite measures on $(X, \mathcal{F})$ such that $\mu_1(X) = \mu_2(X)$ and $\mu_1 = \mu_2$ when restricted to $\mathcal{C}$. Then, $\mu_1 = \mu_2$ on all of $\mathcal{F}$.

↪**Proposition 1.10**: $\{\varnothing\} \cup \{(a, b) : a < b \in \mathbb{R}\}$ a π-system.

↪**Proposition 1.11**: If μ a measure on $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$ such that for all intervals I, $\mu(I) = \ell(I)$, then $\mu = m$.

Proof. Consider for all $n \geq 1$ $\mu|_{\mathfrak{B}_{[-n,n]}}$. Clearly, $\mu([-n, n]) = m([-n, n]) = 2n$, and for all $a, b \in \mathbb{R}$, $\mu((a, b) \cap [-n, n]) = \ell((a, b) \cap [-n, n]) = m((a, b) \cap [-n, n])$. Thus, by the previous theorem, μ must match m on all of $\mathfrak{B}_{[-n,n]}$.

Let now $A \in \mathfrak{B}_{\mathbb{R}}$. Let $A_n := A \cap [-n, n] \in \mathfrak{B}_{[-n,n]}$. By continuity of m from below,

$$\begin{aligned} \mu(A) &= \lim_{n\to\infty} \mu(A_n) \\ &= \lim_{n\to\infty} m(A_n) \\ &= m(A), \end{aligned}$$

hence $\mu = m$. ■

↪**Proposition 1.12**: If μ a measure on $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$ assigning finite values to compact sets and is translation invariant, then $\mu = cm$ for some $c > 0$.

Remark 1.8: This proposition is also tacitly stating that $\mathfrak{B}_{\mathbb{R}}$ translation invariant; this needs to be shown.

↪**Lemma 1.1**: $\mathfrak{B}_{\mathbb{R}}$ translation invariant; for any $A \in \mathfrak{B}_{\mathbb{R}}, x \in \mathbb{R}, A + x \in \mathfrak{B}_{\mathbb{R}}$.

Proof. We employ the "good set strategy"; fix some $x \in \mathbb{R}$ and let

$$\Sigma := \{B \in \mathfrak{B}_{\mathbb{R}} : B + x \in \mathfrak{B}_{\mathbb{R}}\}.$$

We have by construction $\Sigma \subseteq \mathfrak{B}_{\mathbb{R}}$. One can check too that Σ a σ-algebra. But in addition, its easy to see that $\{(a, b) : a < b \in \mathbb{R}\} \subseteq \Sigma$, since a translated interval is just another interval, and since these sets generate $\mathfrak{B}_{\mathbb{R}}$, it must be further that $\mathfrak{B}_{\mathbb{R}} \subseteq \Sigma$, completing the proof. ■

Proof. (of the proposition) Let $c = \mu((0, 1])$, noting that $c > 0$ (why? Consider what would happen if $c = 0$).

This implies that $\forall\, n \geq 1, \mu\left(\left(0, \frac{1}{n}\right]\right) = \frac{c}{n}$ (obtained by "chopping up" $(0, 1]$ into n disjoint intervals); from here we can draw many further conclusions:

$$\begin{aligned} &\forall\, m = 1, ..., n-1, \mu\left(\left(0, \frac{m}{n}\right]\right) = \frac{m}{n} c \\ &\Rightarrow \forall\, q \in \mathbb{Q} \cap (0, 1], \mu((0, q]) = qc \\ &\Rightarrow \forall\, q \in \mathbb{Q}^+, \mu((0, q]) = q \cdot c \text{ (translate)} \\ &\Rightarrow \forall\, a \in \mathbb{R}, \mu((a, a+q]) = q \cdot c \\ &\Rightarrow \forall \text{ intervals } I, \mu(I) = c \cdot \ell(I) \text{ (continuity)} \\ &\Rightarrow \forall\, n \geq 1, a, b \in \mathbb{R}, \mu((a, b) \cap [-n, n]) = c \cdot \ell((a, b) \cap [-n, n]) = c \cdot m((a, b) \cap [-n, n]), \end{aligned}$$

but then, $\mu = c \cdot m$ on $\mathfrak{B}_{\mathbb{R}_{[-n,n]}}$, and by appealing again the Dynkin's, $\mu = c \cdot m$ on all of $\mathfrak{B}_{\mathbb{R}}$. ■

↪**Proposition 1.13** (Scaling): m has the *scaling property* that $\forall\, A \in \mathcal{M}, c \in \mathbb{R}, c \cdot A = \{cx : x \in A\} \in \mathcal{M}$, and $m(c \cdot A) = |c|\, m(A)$.

Proof. Assume $c \neq 0$. Given $A \subseteq \mathbb{R}$, remark that $\{I_n\}$ an open interval cover of A iff $\{cI_n\}$ and open interval cover of cA, and $\ell(cI_n) = |c|\, \ell(I_n)$, and thus $m^*(cA) = |c|\, m^*(A)$.

Now, suppose $A \in \mathcal{M}$. Then, we have for any $B \subseteq \mathbb{R}$,

$$\begin{aligned} m^*(B) = |c|\, m^*\left(\frac{1}{c}B\right) &= |c|\, m^*\left(\frac{1}{c}B \cap A\right) + |c|\, m^*\left(\frac{1}{c}B \cap A^c\right) \\ &= m^*(B \cap cA) + m^*\left(B \cap (cA)^c\right), \end{aligned}$$

so $cA \in \mathcal{M}$. ■

§1.7 Relationship between $\mathfrak{B}_{\mathbb{R}}$ and $\mathcal{M}$

↪**Definition 1.8**: Given $(X, \mathcal{F}, \mu)$, consider the following collection of subsets of X,

$$\mathcal{N} := \{B \subseteq X : \exists\, A \in \mathcal{F} \text{ s.t. } \mu(A) = 0, B \subseteq A\}.$$

Put $\overline{\mathcal{F}} := \sigma(\mathcal{F} \cup \mathcal{N})$; this is called the *completion* of $\mathcal{F}$ with respect to μ.

↪**Proposition 1.14**: $\overline{\mathcal{F}} = \{F \subseteq X : \exists\, E, G \in \mathcal{F} \text{ s.t. } \exists\, E \subseteq F \subseteq G \text{ and } m(G \setminus E) = 0\}$.

Proof. Put $\mathcal{G}$ the set on the right; one can check $\mathcal{G}$ a σ-algebra. Since $\mathcal{F} \subseteq \mathcal{G}$ and $\mathcal{N} \subseteq \mathcal{G}$, we have $\overline{\mathcal{F}} \subseteq \mathcal{G}$.

Conversely, for any $F \in \mathcal{G}$, we have $E, G \in \mathcal{F}$ such that $E \subseteq F \subseteq G$ with $m(G \setminus E) = 0$. We can rewrite

$$F = \underbrace{E}_{\in \mathcal{F}} \cup \underbrace{(F \setminus E)}_{\substack{\subseteq G \setminus E \\ \Rightarrow \mu(F \setminus E) = 0 \\ \Rightarrow G \setminus E \in \mathcal{N}}},$$

hence $F \in \mathcal{F} \cup \mathcal{N}$ and thus in $\overline{\mathcal{F}}$, and equality holds. ■

↪**Definition 1.9**: Given $(X, \mathcal{F}, \mu)$, μ can be *extended* to $\overline{\mathcal{F}}$ by, for each $F \in \overline{\mathcal{F}}$ with $E \subseteq F \subseteq G$ s.t. $\mu(G \setminus E) = 0$, put

$$\mu(F) = \mu(E) = \mu(G).$$

We call then $(X, \overline{\mathcal{F}}, \mu)$ a *complete measure space*.

Remark 1.9: It isn't obvious that this is well defined a priori; in particular, the E, G sets are certainly not guaranteed to be unique in general, so one must check that this definition is valid regardless of choice of "sandwich sets".

↪**Theorem 1.5**: $(\mathbb{R}, \mathcal{M}, m)$ is the completion of $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}}, m)$.

Proof. Given $A \in \mathcal{M}$, then $\forall\, n \geq 1, \exists\, G_n$-open with $A \subseteq G_n$ s.t. $m^*(G_n \setminus A) \leq \frac{1}{n}$ and $\exists\, F_n$-closed with $F_n \subseteq A$ s.t. $m^*(A \setminus F_n) \leq \frac{1}{n}$.

Put $C := \bigcap_{n=1}^{\infty} G_n$, $B := \bigcap_{n=1}^{\infty} F_n$, remarking that $C, B \in \mathfrak{B}_{\mathbb{R}}$, $B \subseteq A \subseteq C$, and moreover

$$m(C \setminus A) \leq \frac{1}{n}, m(A \setminus B) \leq \frac{1}{n}$$

$$\Rightarrow m(C \setminus B) = m(C \setminus A) + m(A \setminus B) \leq \frac{2}{n},$$

but n can be arbitrarily large, hence $m(C \setminus B) = 0$; in short, given a measurable set, we can "sandwich it" arbitrarily closely with Borel sets. Thus, $A \in \overline{\mathfrak{B}_{\mathbb{R}}} \Rightarrow \mathcal{M} \subseteq \overline{\mathfrak{B}_{\mathbb{R}}}$. But recall that $\mathcal{M}$ complete, so $\mathfrak{B}_{\mathbb{R}} \subseteq \mathcal{M} \Rightarrow \overline{\mathfrak{B}_{\mathbb{R}}} \subseteq \overline{\mathcal{M}} = \mathcal{M}$, and thus $\overline{\mathfrak{B}_{\mathbb{R}}} = \mathcal{M}$ indeed.

Heuristically, this means that any measurable set is "different" from a Borel set by at most a null set. ■

§1.8 Some Special Sets

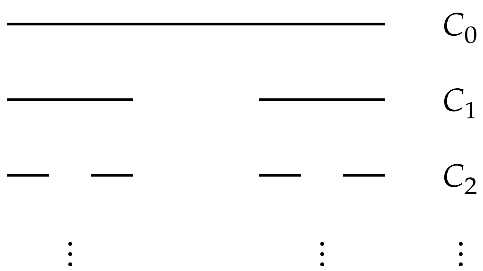
1.8.1 Uncountable Null Set?

Remark that for any countable set $A \in \mathcal{M}$, $m(A) = 0$; indeed, one may write $A = \bigcup_{n=1}^{\infty} \{a_n\}$ for singleton sets $\{a_n\}$, and so

$$m(A) = \sum_{n=1}^{\infty} m(a_n) = 0.$$

One naturally asks the opposite question, does there exist a measurable, *uncountable* set with measure 0? We construct a particular one here, the Cantor set, C.

This requires an "inductive" construction. Define $C_0 = [0, 1]$, and define C_k to be C_{k-1} after removing the middle third from each of its disjoint components. For instance $C_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$, then $C_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right]$, and so on. This may be clearest graphically:



Remark that the $C_n \downarrow$. Put finally

$$C := \bigcap_{n=1}^{\infty} C_n.$$

$\hookrightarrow$**Proposition 1.15**: The following hold for the Cantor set C:

1. C is closed (and thus $C \in \mathfrak{B}_{\mathbb{R}}$);
2. $m(C) = 0$;
3. C is uncountable.

PROOF.

1. For each n, C_n is the countable (indeed, finite) union of 2^n-many disjoint, closed intervals, hence each C_n closed. C is thus a countable intersection of closed sets, and is thus itself closed.
2. For each n, each of the 2^n disjoint closed intervals in C_n has length $\frac{1}{3^n}$, hence
$$m(C_n) = \frac{2^n}{3^n} = \left(\frac{2}{3}\right)^n.$$
Since $\{C_n\} \downarrow$, by continuity of m we have
$$m(C) = \lim_{n \Rightarrow \infty} m(C_n) = \lim_{n \to \infty} \left(\frac{2}{3}\right)^n = 0.$$
3. This part is a little trickier. Notice that for any $x \in [0,1]$, we can define a sequence (a_n) where each $a_n \in \{0,1,2\}$, and such that
$$x = \sum_{n=1}^{\infty} \frac{a_n}{3^n};$$
in particular, this is just the base-3 representation of x, which we denote $(x)_3 = (a_1 a_2 \cdots)$.

I claim now that
$$C = \{x \in [0,1] : (x)_3 \text{ has no 1's}\}.$$
Indeed, at each stage n of the construction of the Cantor set, we get rid of the segment of the real line that would correspond to the $a_n = 1$. One should note that $(x)_3$ not necessarily unique; for instance $\left(\frac{1}{3}\right)_3 = (1,0,0,\ldots) = (0,2,2,\ldots)$, but if we specifically consider all x such that there *exists* a base three representation with no 1′s, i.e. like $\frac{1}{3}$, then C indeed captures all the desired numbers.

Thus, we have that
$$\text{card}(C) = \text{card}(\{\{a_n\} : a_n = 0, 2\}).$$
Define now the function
$$f : C \to [0,1], \quad x \mapsto \sum_{n=1}^{\infty} \frac{a_n}{2} \cdot \frac{1}{2^n}, \text{where } (x)_3 = (a_n)$$
i.e., we "squish" the base-3 representation into a base-2 representation of a number. This is surjective; for any $y \in [0,1]$, $(b_n) := (y)_2$ contains only 0′s and 1′s, hence $(2b_n)$

contains only 0's and 1's, so let x be the number such that $(x)_3 = (2b_n)$. This necessarily exists, indeed, we simply take our definitions backwards:

$$x := \sum_{n=1}^{\infty} \frac{2b_n}{3^n},$$

which maps to y under f and is contained in C. Hence, $\mathrm{card}(C) \geq \mathrm{card}([0,1])$; but $[0,1]$ uncountable, and thus so is C. ■

We can naturally extend the function f used here to map the entire interval $[0,1] \to [0,1]$ as follows

$$f(x) := \begin{cases} \sum_{n=1}^{\infty} \frac{a_n}{2} \cdot \frac{1}{2^n} \text{ if } x \in C, (x)_3 = (a_n) \\ f(a) \text{ if } x \notin C \text{ then } x \in (a,b) \text{ s.t. } (a,b) \text{ removed from } [0,1] \end{cases}.$$

This function is often called the *Devil's Staircase* or *Cantor-Lebesgue function*.

↪**Proposition 1.16**:
1. $f(0) = 0, f(1) = 1, f \equiv \frac{1}{2}$ on $\left(\frac{1}{3}, \frac{2}{3}\right), f \equiv \frac{1}{4}$ on $\left(\frac{1}{9}, \frac{2}{9}\right)$
2. $f : [0,1] \to [0,1]$ a surjection
3. f is nondecreasing
4. f is continuous

Proof. 1., 2., clear from construction.

For 3., let $x_1 < x_2 \in C$, and suppose $(x_1)_3 = (a_n), (x_2)_3 = (b_n)$. Then, since $x_1 < x_2$, it must be that a_n, b_n can only be equal up to some finite N; then the next $0 = a_{N+1} < b_{N+1} = 2$. Hence, it follows that the "modified binary expansion" that arises from f gives directly that $f(x_1) \leq f(x_2)$.

For 4., f is clearly continuous on $[0,1] - C$, since it is piecewise-constant here. Also, f is "one-sided continuous" at each of the "boundary points" $\frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{2}{9}, \ldots$. If $x \in C$, for any $n \geq 1$, there must be x_n, x_n' such that $x_n < x < x_n'$ (if $x = 0$, only need x_n', if $x = 1$, only need x_n) and $f(x_n') - f(x_n) \leq \frac{1}{2^n}$. Then, f is continuous at x by monotonicity of f. ■

1.8.2 Non-Measurable Sets?

We've shown then that there is indeed an uncountable set of measure 0. Another question we may ask ourselves is, is there a $A \subseteq \mathbb{R}$ that is non-measurable? The answer to this turns out to be yes, but the construction requires invoking the axiom of choice:

Axiom 1 (Of Choice): If Σ a collection of nonempty sets, then $\exists$ a function

$$S : \Sigma \to \bigcup_{A \in \Sigma} A,$$

such that $A \in \sigma$, $S(A) \in A$. Such a function is called a *selection function*, and $S(A)$ a *representative* of A.

We construct now a non-measurable set, assuming the above. Consider $[0,1]$, and define an equivalence relation $\sim$ on $[0,1]$ by

$$a \sim b \Leftrightarrow a - b \in \mathbb{Q}.$$

Its easy to check that this is indeed an equivalence relation. Denote by E_a the equivalence class containing a, and set $\Sigma = \{E_a : a \in [0,1]\}$. Note that for any $E_a \in \Sigma, E_a \neq \varnothing$.

Invoking the axiom of choice, we can select exactly one element S_a from E_a for each $E_a \in \Sigma$. Set

$$N := \{S_a : S_a \text{ is a representative of } E_a, E_a \in \Sigma\}.$$

↪**Proposition 1.17**: N, called a *Vitali set*, is non-measurable.

Proof. Assume towards a contradiction that N indeed measurable, $N \in \mathcal{M}$. Consider $[-1,1] \cap \mathbb{Q}$; this is countable, so we can enumerate it $\{q_k\}$, $k \geq 1$. For each k, put

$$N_k := N + q_k.$$

By the assumption of measurability and translation invariance of m, it must be that each N_k measurable and has the same measure as N.

We claim each N_k disjoint. Assume not, then $\exists k \neq \ell$ (i.e. $q_k \neq q_\ell$) and $S_a, S_b \in N$ such that $S_a + q_k = S_b + q_\ell$. But then $S_a - S_b = q_\ell - q_k \in \mathbb{Q}$, hence $S_a \sim S_b$. But we constructed N to have only one representative from each equivalence class, hence it must be that $S_a = S_b$, and so $S_a + q_k = S_a + q_\ell \Rightarrow q_k = q_\ell$, contradicting the assumed distinctness of the q's; hence, the N_k's indeed disjoint.

We claim next that $[0,1] \subseteq \bigcup_{k=1}^{\infty} N_k$. Let $x \in [0,1]$. Then, $x \sim S_a$ for some unique $S_a \in N$ and so $x - S_a \in \mathbb{Q}$. But also, $x, S_a \in [0,1]$, hence $x - S_a \in [-1,1]$ (moreover, $x - S_a \in [-1,1] \cap \mathbb{Q}$) and there must exist a k such that $x - S_a = q_k$, since the q_k's enumerate the entire $[-1,1] \cap \mathbb{Q}$. Thus, $x \in N_k$ by the construction of the N_k's. Thus, $[0,1] \subseteq \bigcup_{n=1}^{\infty} N_k$ indeed.

On the other hand, $\bigcup_{k=1}^{\infty} N_k \subseteq [-1,2]$ and so we have the "bound"

$$[0,1] \subseteq \bigcup_{n=1}^{\infty} N_k \subseteq [-1,2].$$

Taking the measure of all sides then, we have the bound

$$1 \leq \mu\left(\bigcup_{n=1}^{\infty} N_k\right) \leq 3.$$

Invoking the disjointness of the N_k's, we can also use countable additivity to write

$$\mu\left(\bigcup_{n=1}^{\infty} N_k\right) = \sum_{k=1}^{\infty} m(N_k) = \sum_{k=1}^{\infty} m(N),$$

but this final line is a sequence of positive, constant real numbers; hence, it is impossible for it to be within 1 and 3, and we have a contradiction. Hence, N indeed not measurable.

Remark that this proof also shows that $m^*(N_k) > 0$ so $m^*(N) > 0$ (given the interval bound on N we've found). ■

↪**Proposition 1.18**: For every $A \in \mathcal{M}$ such that $m(A) > 0$, there exists $B \subseteq A$ such that B is non-measurable.

Proof. Assume otherwise, that there is a $A \in \mathcal{M}$ with $m(A) > 0$ such that any subset B of A is also measurable.

Remark that $A \subseteq \bigcup_{n \in \mathbb{Z}} A \cap [n, n+1]$. Then, there exists an n such that $m(A \cap [n, n+1]) > 0$ and thus, translating $A' := A \cap [n, n+1] - n)$, $m(A') > 0$, noting that $A' \subseteq [0,1]$. Now, for any $B' \subseteq A'$, $B' + n \subseteq A$. By assumption, then $B' + n$ must be measurable so B' measurable.

In summary, then, we have $A' \subseteq [0,1]$ with $m(A') > 0$ such that (by assumption) B' measurable for all $B' \subseteq A'$.

Let $N, \{q_k\}, N_k$ be as in the previous proof. Set

$$A_k{}' := A' \cap N_k, k \geq 1.$$

Then, $A_k{}'$ disjoint, and

$$A' = [0,1] \cap A' \subseteq \bigcup_{k=1}^{\infty} (N_k \cap A') = \bigcup_{k=1}^{\infty} A_{k'}.$$

Since $m(A') > 0$, there exists a k such that $m(A_k{}') > 0$. Set, for this k,

$$L := \{\ell \geq 1 : q_\ell + q_k \in [-1,1]\}.$$

This set is again countably infinite. We translate, obtaining a disjoint sequence of sets $\{q_\ell + A_k{}' : \ell \in L\}$; since $q_\ell + q_k \in [-1,1] \cap \mathbb{Q}$, then $q_\ell + q_k = q_m$ for some unique m, and so $q_\ell + A_k{}' = q_\ell + A' \cap (N + q_k) \subseteq N_m$. Hence, we have on the one hand that by countable additivity

$$\bigcup_{\ell \in L} (q_\ell + A_k{}') \subseteq [-1,2] \Rightarrow \sum_{\ell \in L} m(q_\ell + A_k{}') \leq 3,$$

and so it must be that $m(q_\ell + A_k{}') = m(A_k{}') = 0$ (else the series couldn't be finite), contradicting the finiteness assumption on $m(A_k{}')$. ■

1.8.3 Non-Borel Measurable Set?

We may ask, is there $A \in \mathcal{M}$ such that $A \notin \mathfrak{B}_{\mathbb{R}}$?

Let $f : [0,1] \to [0,1]$ be the Cantor-Lebesgue function, and put $g(x) = f(x) + x$; note that g is continuous and strictly increasing, and is defined $g : [0,1] \to [0,2]$. Remark that g bijective; the strictly increasing gives injective, and moreover $g(0) = 0, g(1) = 2$ hence by intermediate value theorem it is surjective. Hence, $g^{-1} : [0,2] \to [0,1]$ exists, and is also continuous, so in short g is a homeomorphism; it maps open to open, closed to closed. In particular, if $A \in \mathfrak{B}_{\mathbb{R}}$, then $g(A) \in \mathfrak{B}_{\mathbb{R}}$.

Recall that if (a,b) an open interval that gets removed from the construction of C, then f is constant and so g will map (a,b) to another open interval of the same length $b-a$. Thus,

$$m(g([0,1] \setminus C)) = m([0,1] \setminus C) = 1.$$

Hence, $m(g(C)) = 2 - 1 = 1 > 0$, since $g(C \cup [0,1] \setminus C) = [0,2]$. Hence, there exists a $B \subseteq g(C)$ such that $B \notin \mathcal{M}$, as per the previous proposition.

Let $A := g^{-1}(B)$; then $A \subseteq g^{-1}(g(C)) = C$. Since $m(C) = 0$, $A \in \mathcal{M}$ and $m(A) = 0$. But, $A \notin \mathfrak{B}_{\mathbb{R}}$; if it were, then $g(A) = B \in \mathfrak{B}_{\mathbb{R}}$, since g "maintains" Borel sets, but B is not even Lebesgue measurable and so this is a contradiction).

§2 Integration Theory

§2.1 Measurable Functions

We will be considering functions f defined on $\mathbb{R}$ or some subset of $\mathbb{R}$ that could take positive or negative infinity as its value i.e.

$$f : \mathbb{R} \to \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\},$$

where $\overline{\mathbb{R}}$ the *extended real line*; we say f is $\overline{\mathbb{R}}$-valued. If f never takes $\infty, -\infty$ for any $x \in \mathbb{R}$, we say f finite-valued, or just $\mathbb{R}$-valued.

For all $a \in \mathbb{R}$, we consider inverse images

$$f^{-1}([-\infty, a)) := \{x \in \mathbb{R} : f(x) \in [-\infty, a)\} = \{f < a\},$$

remarking the inclusion of $-\infty$; similarly

$$f^{-1}((a, \infty]) := \{x \in \mathbb{R} : f(x) \in (a, \infty]\} = \{f > a\},$$

and so on, for any $B \subseteq \mathbb{R}$,

$$f^{-1}(B) := \{x \in \mathbb{R} : f(x) \in B\} = \{f \in B\}.$$

Remark that

$$\begin{aligned} f^{-1}(B^c) &= \left(f^{-1}(B)\right)^c \\ f^{-1}(A \cap B) &= f^{-1}(A) \cap f^{-1}(B) \\ f^{-1}(A \cup B) &= f^{-1}(A) \cup f^{-1}(B), \end{aligned}$$

which extend naturally for countable unions/intersections.

↪**Definition 2.1** (Measurable Function): $f : \mathbb{R} \to \overline{\mathbb{R}}$ is *measurable* if $\forall\, a \in \mathbb{R}$,

$$f^{-1}([-\infty, a)) \in \mathcal{M}.$$

↪**Proposition 2.1** (Equivalent Definitions of Measurability):

$$\begin{aligned} f \text{ is measurable} &\Leftrightarrow \forall\, a \in \mathbb{R}, f^{-1}([a,\infty]) \in \mathcal{M} \\ &\Leftrightarrow \forall\, a \in \mathbb{R}, f^{-1}((a,\infty]) \in \mathcal{M} \\ &\Leftrightarrow \forall\, a \in \mathbb{R}, f^{-1}([-\infty,a]) \in \mathcal{M} \end{aligned}$$

Proof. We prove just the last equivalence. Notice that $\forall\, a \in \mathbb{R}$, we can use the commuting of inverse images with countable unions, intersections, complement to write

$$f^{-1}([-\infty,a)) = \bigcup_{n=1}^{\infty} f^{-1}\left(\left[-\infty, a - \frac{1}{n}\right)\right)$$

and

$$f^{-1}([-\infty,a]) = \bigcap_{n=1}^{\infty} f^{-1}\left(\left[-\infty, a + \frac{1}{n}\right)\right).$$

■

↪**Proposition 2.2**: If f finite-valued, Then

$$\begin{aligned} f \text{ is measurable} &\Leftrightarrow \forall\, a < b \in \mathbb{R}, f^{-1}((a,b)) \in \mathcal{M} \\ &\Leftrightarrow \quad \cdots \quad f^{-1}((a,b]) \in \mathcal{M} \\ &\Leftrightarrow \quad \cdots \quad f^{-1}([a,b)) \in \mathcal{M} \\ &\Leftrightarrow \quad \cdots \quad f^{-1}([a,b]) \in \mathcal{M}. \end{aligned}$$

↪**Definition 2.2** (Extended Borel Sigma Algebra): Define the Borel "extended" algebra $\mathfrak{B}_{\overline{\mathbb{R}}}$ of subsets of $\overline{\mathbb{R}}$, defined by

$$\mathfrak{B}_{\overline{\mathbb{R}}} := \sigma(\mathfrak{B}_{\mathbb{R}} \cup \{\{-\infty\},\{\infty\}\}).$$

↪**Proposition 2.3**: $\mathfrak{B}_{\overline{\mathbb{R}}} = \sigma(\{[-\infty,a) : a \in \mathbb{R}\})$.

Proof. For every $a \in \mathbb{R}$, we may write

$$[-\infty, a) = \underbrace{(-\infty, a)}_{\in \mathfrak{B}_{\mathbb{R}}} \cup \{-\infty\} \in \mathfrak{B}_{\overline{\mathbb{R}}},$$

so $\sigma(\{[-\infty, a) : a \in \mathbb{R}\}) \subseteq \mathfrak{B}_{\overline{\mathbb{R}}}$.

Conversely, notice that

$$\{-\infty\} = \bigcap_{n=1}^{\infty} [-\infty, -n),$$

and

$$\{\infty\} = \overline{\mathbb{R}} - \left(\bigcup_{n=1}^{\infty} [-\infty, n) \right),$$

so $\{-\infty\}, \{\infty\} \in \sigma(\{[-\infty, a) : a \in \mathbb{R}\})$. Hence, for any $a \in \mathbb{R}$,

$$(-\infty, a) = [-\infty, a) - \{-\infty\} \in \sigma(\{[-\infty, a) : a \in \mathbb{R}\}),$$

and so $\mathfrak{B}_{\mathbb{R}} \subseteq \sigma(\{[-\infty, a) : a \in \mathbb{R}\})$. $\{-\infty\}, \{\infty\} \in \sigma(\{[-\infty, a) : a \in \mathbb{R}\})$ already, and thus $\mathfrak{B}_{\overline{\mathbb{R}}} \subseteq \sigma(\{[-\infty, a) : a \in \mathbb{R}\})$. ■

↪**Proposition 2.4**: $f : \mathbb{R} \to \overline{\mathbb{R}}$ measurable $\Leftrightarrow$ for all $B \in \mathfrak{B}_{\overline{\mathbb{R}}}, f^{-1}(B) \in \mathcal{M}$.

Proof. $\Leftarrow$ is immediate. For $\Rightarrow$, let $\mathcal{C}$ be a collection of subsets of $\overline{\mathbb{R}}$, then put

$$f^{-1}(\mathcal{C}) := \{f^{-1}(B) : B \in \mathcal{C}\}.$$

By an assignment question (2.6),

$$f^{-1}(\sigma(\mathcal{C})) = \sigma\left(f^{-1}(\mathcal{C})\right).$$

Take $\mathcal{C} = \{[-\infty, a) : a \in \mathbb{R}\}$. Then,

$$f^{-1}(\sigma(\mathcal{C})) = f^{-1}\left(\mathfrak{B}_{\overline{\mathbb{R}}}\right) = \sigma\left(f^{-1}(\{[-\infty, a) : a \in \mathbb{R}\})\right).$$

But f measurable, so $f^{-1}([-\infty, a)) \in \mathcal{M}$ for each $a \in \mathbb{R}$, hence sigma $\left(f^{-1}(\{[-\infty, a) : a \in \mathbb{R}\})\right) \subseteq \mathcal{M}$ and so $f^{-1}(\sigma(\mathcal{C})) \subseteq \mathcal{M}$ completing the proof. ■

↪**Corollary 2.1**: If f finite-valued, then f is measurable $\Leftrightarrow$ for every $B \in \mathfrak{B}_{\mathbb{R}}, f^{-1}(B) \in \mathcal{M}$.

↪**Proposition 2.5**: Given $f : \mathbb{R} \to \overline{\mathbb{R}}$, define the *finite valued component* of f given by

$$f_{\mathbb{R}}(x) := \begin{cases} f(x) : -\infty < f(x) < \infty \\ 0 \text{ otherwise} \end{cases}.$$

Then, f measurable $\Leftrightarrow$ $\forall\, B \in \mathfrak{B}_{\mathbb{R}}, f_{\mathbb{R}}^{-1}(B) \in \mathcal{M}$ AND $\{f = \infty\}, \{f = -\infty\}$ both in $\mathcal{M}$.

Proof. ($\Leftarrow$) For any $a \in \mathbb{R}$,

$$f^{-1}([-\infty, a)) = \{f = -\infty\} \cup f^{-1}((-\infty, a)) = \{f = -\infty\} \cup f_{\mathbb{R}}^{-1}((-\infty, a)),$$

a union of measurable sets and hence is itself measurable.

($\Rightarrow$) Remark that $\{f = \infty\}, \{f = -\infty\} \in \mathcal{M}$ automatically. For any $B \in \mathfrak{B}_{\mathbb{R}}$, we have

$$f_{\mathbb{R}}^{-1}(B) = \{x \in \mathbb{R} : f_{\mathbb{R}}(x) \in B\} = \{x \in \mathbb{R} : f(x) \in B, -\infty < f < \infty\} \cup \{x \in \mathbb{R} : 0 \in B, f(x) = \pm\infty\} \in \mathcal{M}.$$

■

↪**Definition 2.3**: If a statement is true for every $x \in A$ where $A \in \mathcal{M}$ s.t. $m(A^c) = 0$, then we say the statement is true a.e. (almost everywhere).

↪**Proposition 2.6**: If $f : \mathbb{R} \to \overline{\mathbb{R}}$ is measurable and $f = g$ a.e. then g is measurable.

↪**Corollary 2.2**: If f is finite-valued a.e., then f is measurable $\Leftrightarrow f_{\mathbb{R}}$ is measurable $\Leftrightarrow \forall\, a < b \in \mathbb{R}, f^{-1}((a, b)) \in \mathcal{M}$.

↪**Proposition 2.7**: If $f \equiv c$ then f measurable.

If $f = \mathbb{1}_A$ for some $A \subseteq \mathbb{R}$, then f is measurable $\Leftrightarrow A \in \mathcal{M}$.

Proof. Assume $f \equiv c$. Then

$$f^{-1}([-\infty, a)) = \begin{cases} \mathbb{R} \text{ if } c < a \\ \varnothing \text{ if } c \geq a \end{cases} \in \mathcal{M}.$$

Assume now $f = \mathbb{1}_A$. For all $a \in \mathbb{R}$,

$$f^{-1}([-\infty, a)) = \begin{cases} \mathbb{R} \text{ if } a > 1 \\ A^c \text{ if } 0 < a \leq 1 \\ \varnothing \text{ if } a \leq 0 \end{cases} \in \mathcal{M} \Leftrightarrow A \in \mathcal{M}.$$

■

↪**Proposition 2.8**: If f is (finite-valued) continuous, then f is measurable.

Proof. $f : \mathbb{R} \to \mathbb{R}$ continuous $\Leftrightarrow$ for all $G \subseteq \mathbb{R}$ open, $f^{-1}(G)$ open. For all $a < b \in \mathbb{R}$, then $f^{-1}((a, b))$ open so $f^{-1}((a, b)) \in \mathcal{M}$ so f measurable.

In fact, if $f : \mathbb{R} \to \mathbb{R}$ continuous, then for all $B \in \mathfrak{B}_{\mathbb{R}}, f^{-1}(B) \in \mathfrak{B}_{\mathbb{R}}$;

$$f^{-1}(\mathfrak{B}_{\mathbb{R}}) = f^{-1}(\sigma(\{\text{open sets}\})) = \sigma\left(\underbrace{f^{-1}(\{\text{open sets}\})}_{\text{all open}}\right) \subseteq \sigma(\{\text{open sets}\}) = \mathfrak{B}_{\mathbb{R}}.$$

Moreover, if f^{-1} (inverse) exists and is continuous, then for any $B \in \mathfrak{B}_{\mathbb{R}}, f(B) \in \mathfrak{B}_{\mathbb{R}}$. ■

↪**Proposition 2.9**: If $f : \mathbb{R} \to \mathbb{R}$ is measurable and $g : \mathbb{R} \to \mathbb{R}$ is continuous, then $g \circ f$ is measurable.

Remark 2.1: The order matters! The converse doesn't hold in general.

PROOF. For all $a \in \mathbb{R}$,

$$\begin{aligned}(g \circ f)^{-1}((-\infty, a)) &= \{x \in \mathbb{R} : g(f(x)) < a\} \\ &= \left\{x \in \mathbb{R} : f(x) \in g^{-1}([-\infty, a))\right\} \\ &= f^{-1}\left(g^{-1}([-\infty, a))\right) \in \mathcal{M}.\end{aligned}$$

■

↪**Proposition 2.10**: If $f : \mathbb{R} \to \overline{\mathbb{R}}$ is measurable, then:

1. for every $c \in \mathbb{R}$, cf is measurable (in particular $-f$ measurable);
2. $|f|$ is measurable;
3. for every $k \in \mathbb{N}$, f^k is a measurable.

PROOF. We prove just 3. If $k = 0$ this is trivial. For any $a \in \mathbb{R}$,

$$\left(f^k\right)^{-1}([-\infty, a)) = \begin{cases} f^{-1}\left([-\infty, a^{\frac{1}{k}})\right) & \text{if } k \text{ is odd} \\ \varnothing & \text{if } k \text{ is even and } a \leq 0 \\ f^{-1}\left([-a^{\frac{1}{k}}, a^{\frac{1}{k}})\right) & \text{if } k \text{ is even and } a > 0 \end{cases} \in \mathcal{M}.$$

■

↪**Proposition 2.11**: If f, g are two finite-valued measurable functions, then $f + g, f \cdot g, f \vee g := \max\{f, g\}, f \wedge g := \min\{f, g\}$ are measurable functions, where

$$(f \vee g)(x) = \max\{f(x), g(x)\}.$$

PROOF. For all $a \in \mathbb{R}$,

$$\begin{aligned}(f + g)^{-1}([-\infty, a) &= \{x \in \mathbb{R} : f(x) + g(x) < a\} \\ &= \{x \in \mathbb{R} : f(x) < a - g(x)\} \\ &= \bigcup_{q \in \mathbb{Q}} \{x \in \mathbb{R} : f(x) < q < a - g(x)\} \\ &= \bigcup_{q \in \mathbb{Q}} \underbrace{\{x \in \mathbb{R} : f(x) < q\}}_{\in \mathcal{M}} \cap \underbrace{\{x \in \mathbb{R} : g(x) < a - q\}}_{\in \mathcal{M}} \in \mathcal{M}.\end{aligned}$$

This implies, then, that $f - g$ measurable, as are $(f + g)^2$ and $(f - g)^2$, and thus

$$fg = \frac{1}{4}\left[(f+g)^2 - (f-g)^2\right]$$

is measurable.

We have too that

$$f \vee g = \frac{1}{2}(|f-g| + (f+g))$$

and so is measurable, and so

$$f \wedge g = -\max\{-f, -g\} = -(-f \vee -g)$$

is measurable. ■

↪**Corollary 2.3**: If f is measurable, then $f^+ := f \vee 0 = \max\{f, 0\}$ and $f^- := -(f \wedge 0) = \max\{-f, 0\}$ are measurable, as is $f \wedge k$ for any $k \in \mathbb{R}$.

Remark 2.2: Notice that $f = f^+ - f^-$, even with "infinities", and $|f| = f^+ + f^-$.

↪**Proposition 2.12**: Let $\{f_n\}$ be a sequence of measurable functions. Then, $\sup_n f_n$, $\inf_n f_n$, $\limsup_{n\to\infty} f_n$, and $\liminf_{n\to\infty} f_n$ are all measurable (where $(\limsup_{n\to\infty} f_n)(x) := \limsup_{n\to\infty} f_n(x) = \inf_{m\geq 1}\sup_{n\geq m} f_n(x) = \lim_{m\to\infty}\sup_{n\geq m} f_n(x)$).

Proof. To show $\sup_n f_n$ measurable, we will show for all $a \in \mathbb{R}$ $\{\sup_n f_n \leq a\} \in \mathcal{M}$.

$$x \in \left\{\sup_n f_n \leq a\right\} \Leftrightarrow \sup_n f_n(x) \leq a \Leftrightarrow f_n(x) \leq a \,\forall\, n \geq 1 \Leftrightarrow x \in \bigcap_{n=1}^{\infty} \{f_n \leq a\},$$

hence $\{\sup_n f_n \leq a\} = \bigcap_{n=1}^{\infty} \underbrace{\{f_n \leq a\}}_{\in \mathcal{M}} \in \mathcal{M}$ and hence $\sup_n f_n$ is measurable. Note that using $\leq$ was important; $\{\sup_n f_n < a\} \subsetneq \bigcap_{n=1}^{\infty}\{f_n < a\}$, since the $\sup_n f_n$ could equal a. We could say the following, however:

$$\left\{\sup_n f_n < a\right\} = \bigcup_{k=1}^{\infty}\left\{\sup_n f_n \leq a - \frac{1}{k}\right\} = \bigcup_{k=1}^{\infty}\bigcap_{n=1}^{\infty}\left\{f_n \leq a - \frac{1}{k}\right\} \in \mathcal{M}.$$

Next, we have $\inf_n f_n = -\sup_n(-f_n)$ so we are done.

For lim sup, lim inf, we have

$$\limsup_n f_n = \inf_{m\geq 1} \underbrace{\sup_{n\geq m} f_n}_{:= g_m}.$$

g_m is measurable for each $m \geq 1$, hence $\inf_m g_m$ is measurable, hence $\limsup_n f_n$ is measurable. Similar logic follows for lim inf.

We could have show, more directly, that

$$
\begin{aligned}
\left\{\limsup_n f_n < a\right\} &= \left\{\inf_{m\geq 1}\sup_{n\geq m} f_n < a\right\} \\
&= \bigcup_{m=1}^{\infty}\left\{\sup_{n\geq m} f_n < a\right\} \\
&= \bigcup_{m=1}^{\infty}\bigcup_{k=1}^{\infty}\left\{\sup_{n\geq m} f_n \leq a - \frac{1}{k}\right\} \\
&= \bigcup_{m=1}^{\infty}\bigcup_{k=1}^{\infty}\bigcap_{n=m}^{\infty}\left\{f_n \leq a - \frac{1}{k}\right\}.
\end{aligned}
$$

■

↪**Proposition 2.13**: Let $\{f_n\}$ be a sequence of measurable functions. Then, all of the following sets are also measurable:

$$
\begin{aligned}
&\left\{x \in \mathbb{R} : \lim_{n\to\infty} f_n(x) \text{ exists in } \mathbb{R}\right\} =: \left\{\lim_{n\to\infty} f_n \text{ exists in } \mathbb{R}\right\}, \\
&\{\lim f_n = \infty\}, \{\lim f_n = -\infty\}, \{\lim f_n = c \in \mathbb{R}\}.
\end{aligned}
$$

Moreover, if $\lim_{n\to\infty} f_n$ exists (in $\mathbb{R}$ or as $\pm\infty$) a.e. with $f = \lim_{n\to\infty} f_n$ a.e. then f is measurable.

Proof. We have

$$
\begin{aligned}
\{\lim f_n \text{ exists in } \mathbb{R}\} &= \{\limsup f_n = \liminf f_n \text{ and } -\infty < \limsup f_n < \infty\} \\
&= \{-\infty < \liminf f_n < \infty\} \cap \{-\infty < \limsup f_n < \infty\} \cap \{\limsup f_n - \liminf f_n = 0\} \in \mathcal{M}.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\{\lim f_n = c\} &= \left\{x \in \mathbb{R} : \forall\, k \geq 1, \exists\, n \geq 1 \text{ s.t. } \forall\, m \geq n, |f_n(x) - c| \leq \frac{1}{k}\right\} \\
&= \underbrace{\bigcap_{k=1}^{\infty}}_{\forall\, \varepsilon = \frac{1}{k} > 0} \underbrace{\bigcup_{n=1}^{\infty}}_{\exists\, n\geq 1} \underbrace{\bigcap_{m=n}^{\infty}}_{\forall\, m\geq n} \left\{|f_n(x) - c| \leq \frac{1}{k}\right\}.
\end{aligned}
$$

■

§2.2 Approximation by Simple Functions

Given a function $f : \mathbb{R} \to \overline{\mathbb{R}}$, measurable, we may write

$$
f = f^+ - f^-,
$$

where f^+, f^- are non-negative measurable functions; so, it suffices to study non-negative measurable functions. For any $n \geq 1$, we have

$$
f_n^+ := (f^+ \wedge n) \cdot \mathbb{1}_{[-n,n]},
$$

i.e., we cap f^+ at n, and disregard values of f^+ outside of $[-n, n]$; hence we limit our view to a $2n \times n$ "box". Then, f_n^+ is non-negative, measurable, bounded (by n), compactly supported (zero outside a bounded set), and in particular $f_n^+ \uparrow$, with limit

$$\lim_{n\to\infty} f_n^+ = f^+.$$

An identical construction follows for f^- with

$$f_n^- := (f^- \wedge n)\mathbb{1}_{[-n,n]},$$

with $f_n^- \uparrow$ and

$$\lim_{n\to\infty} f_n^- = f^-.$$

Fix some n and consider f_n^+. For $k = 0, 1, 2, ..., 2^n n$, define

$$A_{n,k} := \left\{x \in [-n,n] : \frac{k}{2^n} \leq f_n^+(x) < \frac{k+1}{2^n}\right\} = \left\{\frac{k}{2^n} \leq f_n^+ < \frac{k+1}{2^n}\right\} \cap [-n,n] \in \mathcal{M},$$

noting that $A_{n,k} \cap A_{n,\ell} = \varnothing$ if $k \neq \ell$. Set now

$$\varphi_n := \sum_{k=0}^{n\cdot 2^n} \mathbb{1}_{A_{n,k}} \frac{k}{2^n} = \sum_{k=0}^{n\cdot 2^n} \begin{cases} \frac{k}{2^n} \text{ if in } A_{n,k} \\ 0 \text{ else} \end{cases}.$$

We call φ_n a "simple function"; more generally:

↪**Definition 2.4**: φ is a *simple function* if $\varphi = \sum_{k=1}^{L} \mathbb{1}_{E_k} \cdot a_k$ where L a positive integer, a_k's are constant, E_k's are measurable sets of finite measure.

Moreover, note that $\varphi_n \uparrow$; at each new stage $n \to n+1$, the regions are cut in two, $A_{n,k} = A_{n+1,2k} \cup A_{n+1,2k+1}$. In addition, we have $\varphi_n \leq f_n^+ \leq f^+$ for all n. Moreover, we have the following:

↪**Proposition 2.14**:

$$\lim_{n\to\infty} \varphi_n(x) = f^+(x)$$

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

Proof. For all $x \in \mathbb{R}$, for sufficiently large n we have that $x \in [-n,n]$ and so $f^+(x) = f^+(x)\mathbb{1}_{[-n,n]}(x)$. Assume for now $f^+ < \infty$. Then, for sufficiently large(r?) n, we can ensure $f^+(x) < n$ and so $f^+(x) = f_n^+(x)$ for such an x. Further, we have that $x \in A_{n,k}$ for some k so $\varphi_n(x) = \frac{k}{2^n}$ and $f_n^+(x) < \frac{k+1}{2^n}$ and thus

$$0 \leq f_n^+(x) - \varphi_n(x) < \frac{k+1}{2^n} - \frac{k}{2^n} = 2^{-n}$$

by construction and so $0 \leq f^+(x) - \varphi_n(x) \leq 2^{-n}$ and thus $\lim_{n\to\infty} \varphi_n(x) = f^+(x)$.

In the case that $f^+(x) = \infty$, then $\varphi_n(x) = n$ for all sufficiently large n hence

$$\lim_{n\to\infty} \varphi_n(x) = \lim_{n\to\infty} n = \infty = f^+(x).$$

■

↪**Theorem 2.1**: If g is measurable and non-negative, there exists a sequence of simple functions $\{\varphi_n\}$ such that $\varphi_n \uparrow$ and $\lim_{n\to\infty} \varphi_n(x) = g(x)$ for every $x \in \mathbb{R}$.

We can repeat this same construction and proof for f^- with a sequence $\widetilde{\varphi_n}$. Even better:

↪**Theorem 2.2**: If f is measurable, then $\exists$ a sequence of simple functions $\{\psi_n\}$ such that $|\psi_n| \uparrow$ and $|\psi_n| \leq |f|$ for all n and for all $x \in \mathbb{R}$, $\lim_{n\to\infty} \psi_n(x) = f(x)$.

Proof. Take $\psi_n = \varphi_n - \widetilde{\varphi_n}$ as above; then for all $x \in \mathbb{R}$, at least one of $\varphi_n(x), \widetilde{\varphi_n}(x)$ equals zero. Then

$$|\psi_n| = \varphi_n + \widetilde{\varphi_n} < f^+ + f^- = |f|,$$

and

$$\lim_{n\to\infty} \psi_n(x) = \lim_{n\to\infty} \varphi_n(x) - \lim_{n\to\infty} \widetilde{\varphi_n}(x) = f^+ - f^- = f.$$

■

↪**Definition 2.5** (Step Function): θ a *step function* if it takes the form

$$\theta(x) = \sum_{k=1}^{L} a_k \mathbb{1}_{I_k}(x),$$

where $L \in \mathbb{N}, a_k$'s constant, and I_k finite, open intervals.

↪**Theorem 2.3**: If f is measurable, then there exists a sequence of step functions $\{\theta_n\}$ such that

$$\lim_{n\to\infty} \theta_n(x) = f(x) \text{ for } \textbf{almost every } x \in \mathbb{R}.$$

In particular, we do not have pointwise convergence as for general simple functions, but we have convergence outside a zero-measure set.

Proof. Assume, wlog, that f non-negative (by the previous construction, we can "split" f if not and approximate its positive, negative parts). Given $A \in \mathcal{M}$ with finite measure, recall that for every $\varepsilon > 0$, there exists finitely many finite open intervals $I_1, \ldots, I_N$ such that

$$m\left(A \triangle \left(\bigcup_{i=1}^{N} I_i\right)\right) < \varepsilon.$$

By renaming/rearranging I_i's if necessary, we may assume that I_i's are disjoint; hence

$$\mathbb{1}_{\bigcup_{i=1}^{N} I_i} = \sum_{i=1}^{N} \mathbb{1}_{I_i}.$$

Put

$$\theta_A := \sum_{i=1}^{N} \mathbb{1}_{I_i},$$

noting this is indeed a step function as the name suggests. Then, remark that

$$\underbrace{m(\{x \in \mathbb{R} : \mathbb{1}_A(x) \neq \theta_A(x)\})}_{=A \triangle \left(\bigcup_{n=1}^{N} I_i\right)} < \varepsilon.$$

Since f measurable and non-negative, $\exists \{\varphi_n\}$ sequence of simple functions with limit f. In particular,

$$\varphi_n = \sum_{k=0}^{n2^n} \frac{k}{2^n} \mathbb{1}_{A_{n,k}}.$$

Applying our above analysis to each $A_{n,k}$, then, we have that for any $n \geq 1$ and $k = 0, 1, \ldots, n2^n$ we can find a step function $\theta_{n,k}$ such that

$$m\Big(\Big\{x \in \mathbb{R} : \mathbb{1}_{A_{n,k}} \neq \theta_{n,k}(x)\Big\}\Big) < \frac{1}{2^n(n2^n + 1)} \quad ("= \varepsilon").$$

Put then

$$\theta_n := \sum_{k=0}^{n2^n} \frac{k}{2^n} \theta_{n,k},$$

which is itself a step function. Put

$$E_n := \{x \in \mathbb{R} : \theta_n(x) \neq \varphi_n(x)\}.$$

Then,

$$m(E_n) \leq m\left(\bigcup_{k=0}^{n2^n} \Big\{\theta_{n,k} \neq \mathbb{1}_{A_{n,k}}\Big\}\right) \leq \sum_{k=0}^{n2^n} m\Big(\Big\{\theta_{n,k} \neq \mathbb{1}_{A_{n,k}}\Big\}\Big) \leq 2^{-n}.$$

The φ_n's are chosen such that $\forall\, x \in \mathbb{R}, |\varphi_n(x) - f_n(x)| \leq \frac{1}{2^n}$. Putting

$$F_n := \{x \in \mathbb{R} : |\theta_n(x) - f_n(x)\,| > 2^{-n}\},$$

then remark that $F_n \subseteq E_n$ so $m(F_n) \leq \frac{1}{2^n}$.

We claim now that for a.e. $x \in \mathbb{R}$, $\exists\, m \geq 1$ such that $\forall\, n \geq m, |\theta_n(x) - f_n(x)| \leq \frac{1}{2^n}$, remarking that such an m is *dependent* on x. Consider the complement of this statement; if this set has measure 0, we are done. The logical negation would be "for every $m \geq 1$, exist $n \geq m$ such that $|\theta_n(x) - f_n(x)\,| > 2^{-n}$", which is equivalent to the set

$$\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \{x \in \mathbb{R} : |\theta_n(x) - f_n(x)| > 2^{-n}\} = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} F_n.$$

Let $B_m := \bigcup_{n=m}^{\infty} F_n$; notice $B_m \downarrow$. Then, by continuity from above ****

$$m\left(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}F_n\right) = \lim_{m\to\infty} m(B_m) \leq \lim_{m\to\infty}\sum_{n=m}^{\infty} m(F_n) \leq \lim_{m\to\infty}\sum_{n=m}^{\infty}\frac{1}{2^n} = 0,$$

since the tail of a convergent series must converge to zero. Hence, the set has measure 0 as desired so for almost every $x \in \mathbb{R}$ there exists $m \geq 1$ such that for all $n \geq m$, $|\theta_n - f_n| \leq \frac{1}{2^n}$, hence almost every where $\lim_{n\to\infty}(\theta_n - f_n) = 0$. Therefore, almost everywhere,

$$\theta_n = (\theta_n - f_n) + f_n \overset{n\to\infty}{\longrightarrow} f.$$

■

In this proof, we have proven (and then used) more generally:

↪**Lemma 2.1** (Borel-Cantelli Lemma): If $\{F_n\} \subseteq \mathcal{M}$ such that $\sum_{n=1}^{\infty} m(F_n) < \infty$, then

$$m\left(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}F_n\right) = 0.$$

Proof. Remark that $\bigcup_{n=m}^{\infty}F_n$ a decreasing sequence of functions indexed by m. By continuity of the measure and subadditivity,

$$m\left(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}F_n\right) = \lim_{m\to\infty} m\left(\bigcup_{n=m}^{\infty}F_n\right) \leq \lim_{m\to\infty}\sum_{n=m}^{\infty} m(F_n) = 0,$$

since the tail of a converging sequence must converge to zero. ■

§2.3 Convergence Almost Everywhere vs Convergence in Measure

↪**Definition 2.6** (Convergence Almost Everywhere): For measurable functions $\{f_n\}, f$ we say f_n converges to f a.e. and write $f_n \to f$ a.e. if for almost every $x \in \mathbb{R}$, $\lim_{n\to\infty} f_n(x) = f(x)$.

Similarly, we say $f_n \to f$ a.e. on A if $\exists B \subseteq A$ with $m(B) = 0$ such that $\forall x \in A - B$, $\lim_{n\to\infty} f_n(x) = f(x)$.

↪**Definition 2.7** (Convergence in Measure): For measurable, finite-valued functions $\{f_n\}, f$ we say f_n converges to f in measure and write $f_n \to f$ in measure if for every $\delta > 0$,

$$\lim_{n\to\infty} m(\{x \in \mathbb{R} : |f_n(x) - f(x)| \geq \delta\}) = 0.$$

Similarly, we say $f_n \to f$ in measure on A if $\forall\, \delta > 0$, $\lim_{n\to\infty} m(\{x \in A : |f_n(x) - f(0)| \geq \delta\}) = 0$.

↪**Proposition 2.15**: Given finite-valued measurable functions $\{f_n\}, f$ and $A \in \mathcal{M}$ with finite measure, then if $f_n \to f$ a.e. on A, then $f_n \to f$ in measure on A.

Proof. For all $\delta > 0$,

$$\bigcap_{m=1}^{\infty} \bigcup_{n=m} \{x \in A : |f_n(x) - f(x)| > \delta\} \subseteq \left\{x \in A : \lim_{n\to\infty} f_n(x) \neq f(x)\right\}.$$

The set on the RHS has measure zero and thus so does the left one. Then,

$$\lim_{m\to\infty} m\left(\bigcup_{n=m} \{x \in A : |f_n(x) - f(x)| > \delta\} \right) = 0$$

by continuity, and

$$\{|f_m - f| > \delta\} \subseteq \bigcup_{n=m}^{\infty} \{|f_n - f| > \delta\}$$

hence $m(\{|f_m - f| > \delta\}) \leq m\left(\bigcup_{n=m}^{\infty} \{|f_n - f| > \delta\}\right) \overset{m\to\infty}{\longrightarrow} 0.$ ■

⊛ **Example 2.1**: We give an example of why the assumption that $m(A) < \infty$ is necessary. Let, $f_n = \mathbb{1}_{[n,\infty)}$ and $f \equiv 0$. Then, $\lim_{n\to\infty} f_n(x) = f(x)$ for every $x \in \mathbb{R}$. But $m(\{x \in \mathbb{R} : |f_n(x) - f(x)| = 1\}) = m([n,\infty)) = \infty$.

In general, the converse statement $f_n \to f$ in measure does *not* imply that $f_n \to f$ almost everywhere, even on finite measure sets. Put $\varphi_{1,1} = \mathbb{1}_{[0,1)}$, $\varphi_{2,1} = \mathbb{1}_{\left[0,\frac{1}{2}\right)}$, $\varphi_{2,2} = \mathbb{1}_{\left[\frac{1}{2},1\right)}$, $\varphi_{3,1} = \mathbb{1}_{\left[0,\frac{1}{3}\right)}$, $\varphi_{3,2} = \mathbb{1}_{\left[\frac{1}{3},\frac{2}{3}\right)}$, $\varphi_{3,3} = \mathbb{1}_{\left[\frac{2}{3},1\right)}$, or in general $\varphi_{k,j} = \mathbb{1}_{\left[\frac{j-1}{k},\frac{j}{k}\right)}$ for $j = 1, \ldots, k$. Reorder $\varphi_{k,j}$ "lexicographically" into $\{f_n\}$. Then, we claim $f_n \to 0$ in measure on $[0,1)$; for any $\delta \in (0,1)$,

$$m(\{|f_n - 0| > \delta\}) = \frac{1}{k(n)} \to 0,$$

where $k(n)$ the "row" that f_n comes from. Hence, f_n converges in measure. However, f_n does not converge almost eveywhere on $[0,1)$. Indeed, for each $x \in \mathbb{R}$ and $k \geq 1$, there exists a *unique* j such that $x \in \left[\frac{j-1}{k}, \frac{j}{k}\right]$ hence $\varphi_{k,j}(x) = 1$, so in other notation there always exists an n such that $f_n(x) = 1$, and so precisely $f_n(x) = 1$ for infinitely many n. Hence, we do not have convergence everywhere (in fact, anywhere).

↪**Proposition 2.16**: Given $\{f_n\}, f$ measurable, finite-valued functions, if $f_n \to f$ in measure, then there exists a subsequence $\left\{f_{n_k}\right\}$ such that $f_{n_k} \to f$ a.e. as $k \to \infty$.

Proof. Assume $f_n \to f$ in measure, that is for every $\delta > 0$, $m(\{|f_n - f| > \delta\}) \to 0$. Hence, for all $k \geq 1$, with $\delta = \frac{1}{k}$, we have that for some sufficiently large n_k, we have that $m\left(\underbrace{\left\{|f_{n_k} - f| > \frac{1}{k}\right\}}_{:=A_k}\right) \leq \frac{1}{k^2}$, hence $\sum_{k=1}^{\infty} m(A_k) < \infty$. Hence,

$$m\left(\bigcap_{\ell=1}^{\infty} \bigcup_{k=\ell}^{\infty} A_k\right) = \lim_{\ell\to\infty} m\left(\bigcup_{k=\ell}^{\infty} A_k\right) \leq \lim_{\ell\to\infty} \sum_{k=\ell}^{\infty} m(A_k) = 0,$$

since $\sum_{k=\ell}^{\infty} m(A_k)$ the tail of a converging series. Hence, complementing the above, a.e. there $\exists\, \ell$ such that for every $k \geq \ell$, $|f_{n_k} - f| \leq \frac{1}{k}$ and so $\lim_{k\to\infty} |f_{n_k} - f| = 0$ almost everywhere, and so $f_{n_k} \to f$ a.e. (as $k \to \infty$). ∎

↪**Proposition 2.17** (Subsequence Test): Given $\{f_n\}, f$ measurable, finite-valued functions, $f_n \to f$ in measure $\Leftrightarrow$ for every subsequence $\{n_k\}$, there exists a subsubsequence $\left\{n_{k_\ell}\right\} \subset \{n_k\}$ such that $f_{n_{k_\ell}} \to f$ in measure as $\ell \to \infty$.

Proof. $\Rightarrow$ is clear. For $\Leftarrow$, suppose towards a contradiction that $f_n \nrightarrow f$ in measure. Then, $\exists\, \delta > 0$ and subsequence $\{n_k\}$ $m\left(\left\{|f_{n_k} - f| > \delta\right\}\right) > \delta$ for every k. By the assumption of the RHS, there exists a further subsequence $\left\{n_{k_\ell}\right\}$ such that $f_{n_{k_\ell}} \to f$ in measure. This is a contradiction. ∎

⊛ **Example 2.2** (Assignment Exercise): Prove that if $f_n \to f$ in measure and $g_n \to g$ in measure, $f_n g_n \to fg$ in measure (everything finite valued, measurable).

§2.4 Egorov's Theorem and Lusin's Theorem

Recall that if f is measurable, then $\exists\, \{\theta_n\}$ sequence of step functions such that $\theta_n \to f$ almost everywhere.

↪**Theorem 2.4** (Egorov's): Given $\{f_n\}, f$ measurable functions and $A \in \mathcal{M}$ with $m(A) < \infty$, if $f_n \to f$ a.e. on A, then $\forall\, \varepsilon > 0$, there exists a closed subset $A_\varepsilon \subseteq A$ with $m(A \setminus A_\varepsilon) \leq \varepsilon$ such that $f_n \to f$ uniformly on A_ε.

Proof. We assume first f is finite-valued on A (otherwise, replace A with $A \cap \{-\infty < f < \infty\}$; we'll deal with $\{f = \pm\infty\}$ later). We want to show that $\forall\, \varepsilon > 0, \exists$ closed $A_\varepsilon \subseteq A$ s.t. $m(A \setminus A_\varepsilon) < \varepsilon$ *and* $\sup_{x \in A_\varepsilon} |f_n(x) - f(x)| \to 0$ as $n \to \infty$.

For each $k \geq 1$ and $n \geq 1$, put

$$E_n^{(k)} := \left\{x \in A : |f_j(x) - f(x)| \leq \frac{1}{k}\ \forall\, j \geq n\right\}.$$

For fixed k, remark that $E_n^{(k)} \subseteq E_{n+1}^{(k)}$, i.e. $E_n^{(k)}$ increasing (wrt n), so we may consider

$$\bigcup_{n=1}^{\infty} E_n^{(k)} = \left\{x \in A : \exists\, n \geq 1 \text{ s.t. } \forall\, j \geq n, |f_j(x) - f(x)| \leq \frac{1}{k}\right\} \supseteq \left\{x \in A : \lim_{n\to\infty} f_n(x) = f(x)\right\} =: A'.$$

By assumption, $m(A') = m(A)$, so by continuity and the superset relation above, $m(A) = m(A') \leq m\left(\bigcup_{n=1}^{\infty} E_n^{(k)}\right) = \lim_{n\to\infty} m\left(E_n^{(k)}\right) \leq m(A)$, and thus $\lim_{n\to\infty} m\left(E_n^{(k)}\right) = m(A)$ for every $k \geq 1$.

Given, then, any $\varepsilon > 0$, there exists a n_k such that $m\left(A \setminus E_{n_k}^{(k)}\right) = m(A) - m\left(E_{n_k}^{(k)}\right) < \frac{1}{2^k}\frac{\varepsilon}{2}$. Set

$$B := A \setminus \left(\bigcap_{k=1}^{\infty} E_{n_k}^{(k)} \right),$$

then

$$m(B) = m\left(\bigcup_{k=1}^{\infty} A \setminus E_{n_k}^{(k)} \right) \leq \sum_{k=1}^{\infty} m\left(A \setminus E_{n_k}^{(k)}\right) \leq \frac{\varepsilon}{2}.$$

Put

$$\tilde{A} := A \setminus B = \bigcap_{k=1}^{\infty} E_{n_k}^{(k)}.$$

Then, if $x \in \tilde{A}$, then $x \in E_{n_k}^{(k)}$ for every k, and hence for every $k \geq 1$ and $j \geq n_k$, $|f_j(x) - f(x)| \leq \frac{1}{k}$. This shows then that $f_n \to f$ uniformly on $\tilde{A}$. By regularity of m, there exists a closed $A_\varepsilon \subseteq \tilde{A}$ such that $m(\tilde{A} \setminus A_\varepsilon) \leq \frac{\varepsilon}{2}$. Then, $f_n \to f$ uniformly on A_ε, and $m(A \setminus A_\varepsilon) = m(A \setminus \tilde{A}) + m(\tilde{A} \setminus A_\varepsilon) < \varepsilon$.

Now, if $f = \infty / -\infty$ on A, then $A = A^{\infty} \cup A^{-\infty} \cup A^{\mathbb{R}}$ (with $A^{\bullet} := \{f = \bullet\} \cap A$). The last case is done. For A^{∞} (similar construction for $A^{-\infty}$), define for every $k, n \geq 1$,

$$E_n^{(k)} := \left\{x \in A : f_j(x) > k \; \forall \, j \geq n\right\}.$$

Then, the remainder of the proof follows precisely the same for the sequence of sets $E_n^{(k)}$. ■

Remark 2.3:

1. The assumption $m(A) < \infty$ is necessary. For instance $f_n = \mathbb{1}_{[n,\infty)} \to 0$ pointwise, but for any $a \in \mathbb{R}$, f_n does not converge to 0 uniformly on (a, ∞).
2. In general, Egorov's $\not\Rightarrow$ $f_n \to f$ uniformly a.e.. For instance, on $[0,1]$, let $f_n(x) = x^n$ and $f(x) \equiv 0$. For every $x \in [0,1)$, $f_n(x) \to f(x)$ as $n \to \infty$. Hence, $f_n \to f$ a.e. on $[0,1]$ (the only point that doesn't converge, indeed, is at 1). If $A \subseteq [0,1]$ is closed such that $1 \in A$, then $f_n \nrightarrow f$ uniformly on A. To see this, let $\{x_m\} \subseteq A$ such that $x_m \uparrow$ and $\lim_{m\to\infty} x_m = 1$. Then, for any fixed n,

$$\sup_{x \in A} |f_n(x) - f(x)| \geq \sup_m |f_n(x_m) - f(x_m)| = \sup_m x_m^n = 1,$$

hence f_n does not converge uniformly on A.

↪**Theorem 2.5** (Lusin's Theorem): Given f measurable and finite-valued and $A \in \mathcal{M}$ with $m(A) < \infty$, for all $\varepsilon > 0$, there exists a closed $A_\varepsilon \subseteq A$ with $m(A \setminus A_\varepsilon) < \varepsilon$ such that $f|_{A_\varepsilon}$ is continuous.

Remark 2.4: Lusin's Theorem states that $f|_{A_\varepsilon}$ is continuous as a function on A_ε, which is *not* the same as saying f as a function on A is continuous at points in A_ε.

For instance, $f = \mathbb{1}_{\mathbb{Q}\cap[0,1]}$ is not continuous anywhere on $[0,1]$. However, $f|_{\mathbb{Q}\cap[0,1]}$ is constant and therefore continuous *on* $\mathbb{Q} \cap [0,1]$.

Proof. Let $\{\theta_n\}$ be a sequence of step functions such that $\theta_n \to f$ a.e. on A. Note that θ_n piecewise constant and hence piecewise continuous. Given $\varepsilon > 0$ and $n \geq 1$, we can find an open set E_n such that $\theta_n|_{E_n^c}$ is continuous and $m(E_n) \leq \frac{\varepsilon}{2}\frac{1}{2^n}$. Meanwhile, Egorov's implies that there exists a closed $B \subseteq A$ such that $m(A \setminus B) \leq \frac{\varepsilon}{2}$ such that $\theta_n \to f$ uniformly on B. Set

$$A_\varepsilon = B \setminus \bigcup_{n=1}^{\infty} E_n,$$

noting that $A_\varepsilon \subset A$ closed and

$$m(A \setminus A_\varepsilon) = m(A \setminus B) + m\left(\bigcup_{n=1}^{\infty} E_n\right) = \frac{\varepsilon}{2} + \sum_{n=1}^{\infty} m(E_m) \leq \varepsilon.$$

Finally, on A_ε, $\theta_n \to f$ uniformly and $\theta_n|_{A_\varepsilon}$ continuous, and hence $f|_{A_\varepsilon}$ continuous (uniform limit of continuous functions is continuous). ■

Remark 2.5:

1. Lusin's Theorem $\not\Rightarrow f$ is continuous almost everywhere in general. For instance, recall that fat Cantor set $\tilde{C}$, with $m(\tilde{C}) = \frac{1}{2}$. Let $f = \mathbb{1}_{\tilde{C}}$. f is NOT continuous a.e. on $[0,1]$, i.e. $\forall\, B \subseteq [0,1]$ with $m(B) = 1$, $f|_B$ is NOT continuous. To see this, let $\tilde{D} = [0,1] \setminus \tilde{C}$. Since $m(B) = 1$, then $m(\tilde{C} \cap B) = m(\tilde{D} \cap B) = \frac{1}{2}$. Then for any $x \in \tilde{C} \cap B$, $f|_B$ is NOT continuous at x. If it were at say some $x_0 \in \tilde{C} \cap B$, then there must exist some $\delta > 0$ such that for any $x \in (x_0 - \delta, x_0 + \delta) \cap B$, $|f(x) - f(x_0)| < \frac{1}{2}$. Hence, for any $x \in (x_0 - \delta, x_0 + \delta) \cap B$, $\frac{1}{2} \leq f(x) \leq \frac{3}{2}$. However, $m((x_0 - \delta, x_0 + \delta) \cap B \cap \tilde{D}) > 0$ so it must be that $\exists\, y \in (x_0 - \delta, x_0 + \delta) \cap B \cap \tilde{D} \Rightarrow f(y) = 0$, a contradiction. How, then, does one apply Lusin's; that is, $\forall\, \varepsilon > 0$, there must exist some $A_\varepsilon \subseteq [0,1]$ such that $m([0,1] \setminus A_\varepsilon) < \varepsilon$ and $f|_{A_\varepsilon} < \varepsilon$ (exercise)?
2. (Exercise) The $\{\theta_n\}$'s are not continuous on $\mathbb{R}$, but you can choose a sequence $\left\{\widetilde{\theta_n}\right\}$ to be continuous on $\mathbb{R}$ such that $\widetilde{\theta_n} \to f$ a.e..
3. Lusin's Theorem $\Rightarrow\ \forall\, k$ sufficiently large, $\exists\, A_k \subseteq A$ closed such that $m(A \setminus A_k) \leq \frac{1}{k}$ and $f|_{A_k}$ continuous on A_k. In fact, we can construct them such that $A_k \uparrow$ (otherwise replace A_k with $\bigcup_{i=1}^{k} A_i$).

§2.5 Construction of Integrals

2.5.1 Integral of Simple Functions

↪**Definition 2.8**: Given a simple function $\varphi = \sum_{k=1}^{L} a_k \mathbb{1}_{E_k}$, the *(Lebesgue) integral* of φ is defined as

$$\int_{\mathbb{R}} \varphi(x)\, \mathrm{d}x = \int_{\mathbb{R}} \varphi := \sum_{k=1}^{L} a_k \cdot m(E_k).$$

For any $A \in \mathcal{M}$, $\mathbb{1}_A \varphi$ is again a simple function and we define

$$\int_A \varphi := \int_{\mathbb{R}} \mathbb{1}_A \varphi.$$

↪**Proposition 2.18** (Properties of $\int_{\mathbb{R}} \varphi$):

1. (Well-definedness) The written representation of φ is not necessarily unique, but if $\varphi = \sum_{k=1}^{L} a_k \mathbb{1}_{E_k} = \sum_{\ell=1}^{M} b_\ell \mathbb{1}_{F_\ell}$, then
$$\sum_{k=1}^{L} a_k m(E_k) = \sum_{\ell=1}^{M} b_\ell m(F_\ell).$$
2. (Linearity) If φ, ψ two simple functions and $a, b \in \mathbb{R}$, then $a\varphi + b\psi$ a simple function, and
$$\int_{\mathbb{R}} a\varphi + b\psi = a \cdot \int_{\mathbb{R}} \varphi + b \cdot \int_{\mathbb{R}} \psi.$$
3. (Finite Additivity) If φ a simple function, $A, B \in \mathcal{M}$ with $A \cap B = \varnothing$, then
$$\int_{A \cup B} \varphi = \int_A \varphi + \int_B \varphi.$$
4. (Monotonicity) If φ, ψ are two simple functions with $\varphi \leq \psi$, then $\int_{\mathbb{R}} \varphi \leq \int_{\mathbb{R}} \psi$.
5. If φ a simple function then so is $|\varphi|$ and $|\int_{\mathbb{R}} \varphi| \leq \int_{\mathbb{R}} |\varphi|$.

Proof.

1. wlog, we may assume E_k and F_ℓ are respectively disjoint. Set $a_0 = b_0 = 0$, $E_0 := \left(\bigcup_{k=1}^{L} E_k\right)^c, F_0 := \left(\bigcup_{\ell=1}^{M} F_\ell\right)^c$ for convenience. Now, $\{E_0, ..., E_L\}, \{F_0, ..., F_M\}$ are two partitions of $\mathbb{R}$. In particular, then, for each k, $\mathbb{1}_{E_k} = \sum_{\ell=0}^{M} \mathbb{1}_{E_k \cap F_\ell}$, since $E_k = \bigsqcup_{\ell=0}^{M} (E_k \cap F_\ell)$. Now, we have
$$\varphi = \sum_{k=0}^{L} a_k \mathbb{1}_{E_k} = \sum_{k=0}^{L} \sum_{\ell=0}^{M} a_k \mathbb{1}_{E_k \cap F_\ell}.$$
Similarly partitioning, we have
$$\varphi = \sum_{\ell=0}^{M} b_\ell \mathbb{1}_{F_\ell} = \sum_{\ell=0}^{M} \sum_{k=0}^{L} b_\ell \mathbb{1}_{E_k \cap F_\ell}.$$
If $E_k \cap F_\ell \neq \varnothing$, then $a_k = b_\ell$, and thus on the one hand
$$\int_{\mathbb{R}} \varphi = \sum_{k=0}^{L} \sum_{\ell=0}^{M} a_k m(E_k \cap F_\ell)$$
and on the other

$$\int_{\mathbb{R}} \varphi = \sum_{\ell=0}^{M} \sum_{k=0}^{L} b_\ell m(E_k \cap F_\ell),$$

(with summation convention $0 \cdot \infty = 0$). If $m(E_k \cap F_\ell) > 0$, then $E_k \cap F_\ell \neq \varnothing$ and so $a_k = b_\ell$ and so the two sums agree.

4. Assume $\varphi = \sum_{k=1}^{L} a_k \mathbb{1}_{E_k}, \psi = \sum_{\ell=1}^{M} b_\ell \mathbb{1}_{F_\ell}$. Repeat the partitioning/rewriting steps from part 1, then note that since $\varphi \leq \psi$, if $E_k \cap F_\ell \neq \varnothing$, it must be that $a_k \leq b_\ell$, so if $m(E_k \cap F_\ell) > 0$ $a_k \leq b_\ell$ and thus the monotonicity follows.

■

2.5.2 Integral of Non-Negative Functions

↪**Definition 2.9**: If f a non-negative, measurable function then the integral of f is given by

$$\int_{\mathbb{R}} f(x) \, \mathrm{d}x = \int_{\mathbb{R}} f := \sup \left\{ \int_{\mathbb{R}} \varphi : \varphi \text{ is simple and } \varphi \leq f \right\}.$$

↪**Proposition 2.19**: The definition above agrees with that for simple functions that are also non-negative, namely this definition is consistent with the previous.

Proof. Let φ be non-negative. Then $\varphi \leq \varphi$ certainly so the first definition $\int_{\mathbb{R}} \varphi \leq \sup\{\cdots\}$. Conversely, it suffices to show that for any non-negative simple $\psi \leq \varphi$, $\int_{\mathbb{R}} \psi \leq \int_{\mathbb{R}} \varphi$, using the first definition. But this simply follows from monotonicity of $\int$, and we are done. ■

Remark 2.6: Given $f \geq 0$ and measurable, this definition implies that there exists a sequence $\{\varphi_n\}$ of simple functions such that $\varphi_n \leq f$ and $\lim_{n \to \infty} \int_{\mathbb{R}} \varphi_n = \int_{\mathbb{R}} f$. We would like to show that, in some sense, the choice of $\{\varphi_n\}$ is arbitrary.

↪**Theorem 2.6**: Suppose $f \geq 0$ and measurable. If $\{\varphi_n\}$ a sequence of simple functions such that $\varphi_n \uparrow$ and $\lim_{n \to \infty} \varphi_n = f$ pointwise, then

$$\lim_{n \to \infty} \int_{\mathbb{R}} \varphi_n = \int_{\mathbb{R}} f.$$

Proof. Since $\varphi_n \leq f$ for all $n \geq 1$, then $\int_{\mathbb{R}} \varphi_n \leq \int_{\mathbb{R}} f$ and so $\lim_{n \to \infty} \int_{\mathbb{R}} \varphi_n \leq \int_{\mathbb{R}} f$ (nothing the limit on the LHS necessarily always exists by monotonicity). On the other hand, it suffices to show that $\forall \psi \leq f$ simple, that $\int_{\mathbb{R}} \psi \leq \lim_{n \to \infty} \int_{\mathbb{R}} \varphi_n$. Assume $\psi = \sum_{k=1}^{L} a_k \mathbb{1}_{E_k} = \sum_{k=0}^{L} a_k \mathbb{1}_{E_k}$ where $\{E_0, ..., E_L\}$ forms a partition of $\mathbb{R}$. Since

$$\int_{\mathbb{R}} \psi = \sum_{k=0}^{L} a_k m(E_k)$$

and

$$\int_{\mathbb{R}} \varphi_n = \sum_{k=0}^{L} \int_{E_k} \varphi_n$$

by finite additivity. It suffices to show then that for each $k = 0, \dots, L$, $a_k m(E_k) \leq \lim_{n\to\infty} \int_{E_k} \varphi_n$.

First, if $a_k = 0$ or $m(E_k) = 0$, then we are done. Assume $a_k, m(E_k) > 0$. For each fixed k, $\lim_{n\to\infty} \varphi_n = f \geq \psi$ so for every $x \in E_k$, $\lim_{n\to\infty} \varphi_n(x) \geq \psi(x) = a_k$. For any $\varepsilon > 0$, put

$$C_n^{\varepsilon} := \{x \in E_k : \varphi_n(x) \geq (1-\varepsilon)a_k\}.$$

Since $\varphi_n \leq \varphi_{n+1}$, $C_n^{\varepsilon} \uparrow$ wrt n. Then note

$$\bigcup_{n=1}^{\infty} C_n^{\varepsilon} = E_k.$$

Then,

$$\lim_{n\to\infty} \int_{E_k} \varphi_n = \lim_{n\to\infty} \int_{\mathbb{R}} \mathbb{1}_{E_k}\varphi_n \geq \lim_{n\to\infty} \int_{\mathbb{R}} \mathbb{1}_{C_n^{\varepsilon}}\varphi_n \geq \lim_{n\to\infty} (1-\varepsilon)a_k m(C_n^{\varepsilon}) = (1-\varepsilon)a_k m(E_k),$$

where we use the fact that $\mathbb{1}_{E_k}\varphi_n \geq \mathbb{1}_{C_n^{\varepsilon}}\varphi_n \geq (1-\varepsilon)a_k \mathbb{1}_{C_k^{\varepsilon}}$ and $\lim_{n\to\infty} m(C_n^{\varepsilon}) = m\left(\bigcup_{n=1}^{\infty} C_n^{\varepsilon}\right) = m(E_k)$. Since ε arbitrary, then

$$\lim_{n\to\infty} \int_{E_k} \varphi_n \geq a_k m(E_k),$$

and we are done. ∎

↪**Corollary 2.4**: For any $f \geq 0$ measurable, if $\forall\, n \geq 1, k = 0, 1, \dots, n2^n$ with $A_{n,k} := \left\{\frac{k}{2^n} \leq f < \frac{k+1}{2^n}\right\}$, then

$$\int_{\mathbb{R}} f = \lim_{n\to\infty} \sum_{k=0}^{n2^n} \frac{k}{2^n} m(A_{n,k}).$$

Proof. Let $\varphi_n = \sum_{k=0}^{n2^n} \frac{k}{2^n} \mathbb{1}_{A_{n,k}}$, then $\varphi_n \uparrow$ and $\varphi_n \to f$. ∎

↪**Proposition 2.20** (Properties of Integral of Non-Negative Functions):

1. (Well-definedness) If $f, g \geq 0$ measurable such that $f = g$ a.e., then $\int_{\mathbb{R}} f = \int_{\mathbb{R}} g$.
2. (Linearity) For any $f, g \geq 0$ measurable and $a, b \geq 0$, then $\int_{\mathbb{R}} (af + bg) = a\int_{\mathbb{R}} f + b\int_{\mathbb{R}} g$.
3. (Monotonicity) If $f, g \geq 0$ measurable and $f \leq g$ a.e., then $\int_{\mathbb{R}} f \leq \int_{\mathbb{R}} g$.
4. i. Let $f \geq 0$ measurable, then $\int_{\mathbb{R}} f = 0 \Leftrightarrow f \equiv 0$ a.e.
   ii. Let $f \geq 0$ measurable, $A \in \mathcal{M}$. Then $\int_A f = 0 \Leftrightarrow$ either $f \equiv 0$ a.e. on A or $m(A) = 0$.
   iii. Let $f \geq 0$ measurable, then if $\int_{\mathbb{R}} f < \infty$ then f is finite valued a.e.
5. (Markov's Inequality) Let $f \geq 0$ measurable and $0 < a < \infty$. Then, $m(\{f > a\}) \leq \frac{1}{a}\int_{\mathbb{R}} f$. In particular, if the RHS is finite, $\lim_{\{a\to\infty\}} m(\{f > a\}) = 0$, in fact in $\mathcal{O}\left(\frac{1}{a}\right)$.

Proof.

1. Let $\{\varphi_n\}, \{\psi_n\}$ sequences of simple functions such that both are monotonically increasing with $\varphi_n \to f, \psi_n \to g$. Put $h_n := \varphi_n \mathbb{1}_{\{f=g\}} + \psi_n \mathbb{1}_{\{f \neq g\}}$; then h_n again simple, $h_n \uparrow$, and $h_n \to g$ everywhere. Then,

$$\int_{\mathbb{R}} g = \lim_n \int_{\mathbb{R}} h_n = \lim_n \left(\int_{\{f=g\}} \varphi_n + \int_{\{f \neq g\}} \psi_n \right) = \lim_n \int_{\{f=g\}} \varphi_n.$$

Meanwhile,

$$\int_{\mathbb{R}} f = \lim_n \int_{\mathbb{R}} \varphi_n = \lim_n \left(\int_{\{f=g\}} \varphi_n + \int_{\{f \neq g\}} \varphi_n \right) = \lim_n \int_{\{f=g\}} \varphi_n,$$

and so $\int_{\mathbb{R}} f = \int_{\mathbb{R}} g$.

2. Take $\{\varphi_n\}, \{\psi_n\}$ as in the previous proof. Then $\{h_n : a\varphi_n + b\psi_n\}$ again a sequence of monotonically increasing simple functions with limit $af + bg$. Then

$$\int_{\mathbb{R}} (af + bg) = \lim_n \int_{\mathbb{R}} h_n = \lim_n \int_{\mathbb{R}} (a\varphi_n + b\psi_n) = \lim_n \left(a \int_{\mathbb{R}} \varphi_n + b \int_{\mathbb{R}} \psi_n \right) = a \int_{\mathbb{R}} f + b \int_{\mathbb{R}} g.$$

3. wlog, assume that $f \leq g$ everywhere by replacing f with $f \mathbb{1}_{\{f \leq g\}}$. Then, $\{\varphi :$ simple, $\varphi \leq f\} \subseteq \{\varphi :$ simple, $\varphi \leq g\}$ and so $\int_{\mathbb{R}} f \leq \int_{\mathbb{R}} g$.

4. i. $\Leftarrow$ clear. Conversely, we would like to prove that if $A = \{f > 0\}, m(A) = 0$. Put $A_n := \left\{f \geq \frac{1}{n}\right\}$ for $n \geq 1$. Then, $A_n \uparrow$ and $\bigcup_{n=1}^{\infty} A_n = A$. By continuity of m,

$$m(A) = \lim_n m(A_n).$$

Suppose towards a contradiction that $m(A) = \delta > 0$. Then, $\delta = \lim_n m(A_n)$, and so must exist $N \geq 1$ such that $m(A_N) \geq \frac{\delta}{2}$. Since $f \geq f \mathbb{1}_{A_N} \geq \frac{1}{N} \mathbb{1}_{A_N}$. By monotonicity, $\int_{\mathbb{R}} f \geq \int_{\mathbb{R}} \frac{1}{N} \mathbb{1}_{A_N} = \frac{1}{N} m(A_N) \geq \frac{1}{N} \frac{\delta}{2} > 0$, a contradiction.
ii. By i., $\int_A f = 0 \Leftrightarrow \mathbb{1}_A f \equiv 0$ a.e. on $\mathbb{R}$. If $m(A) = 0$, then $\mathbb{1}_A \equiv 0$ a.e. so $\mathbb{1}_A f \equiv 0$ a.e.. Else, if $m(A) > 0$, then $f \equiv 0$ a.e. on A.
iii. Put $A := \{f = \infty\}$. Assume towards a contradiction that $m(A) = \delta > 0$. Then, for every $n \geq 1, f \geq f \mathbb{1}_A \geq n \mathbb{1}_A$ and so $\int_{\mathbb{R}} f \geq \int_{\mathbb{R}} n \mathbb{1}_A = nm(A) = n\delta$. But this holds for any arbitrary n, so $\int_{\mathbb{R}} f = \infty$, a contradiction.
5. Put $A_a := \{f > a\}$. Then $f \geq f \mathbb{1}_{A_a} > a \mathbb{1}_{A_a}$ so $\int_{\mathbb{R}} f \geq am(A_a)$.

■

2.5.3 Integral of General Measurable, Integrable Functions

↪**Definition 2.10**: For f measurable, $\int_{\mathbb{R}} f := \int_{\mathbb{R}} f^+ - \int_{\mathbb{R}} f^-$, provided that at least one of $\int_{\mathbb{R}} f^+$, $\int_{\mathbb{R}} f^-$ is finite; in particular, $\int_{\mathbb{R}} f$ may be finite or infinite.

Remark 2.7: Only having $\int_{\mathbb{R}} f$ being defined is not sufficient for the desirable properties (linearity, monotonicity) to hold.

↪**Definition 2.11** (Integrable): A measurable function f is called *integrable*, denoted $f \in L^1(\mathbb{R})$, if *both* $\int_{\mathbb{R}} f^+ < \infty$ and $\int_{\mathbb{R}} f^- < \infty$. Note that

$$f \in L^1(\mathbb{R}) \Leftrightarrow \int_{\mathbb{R}} |f| < \infty \ (\text{since} \int_{\mathbb{R}} |f| = \int_{\mathbb{R}} f^+ + \int_{\mathbb{R}} f^-)$$
$$\Leftrightarrow \int_{\mathbb{R}} f \text{ finite valued.}$$

↪**Proposition 2.21** (Properties of Integrals of Integrable Functions):

1. $|\int_{\mathbb{R}} f| \leq \int_{\mathbb{R}} |f|$
2. $f \in L^1(\mathbb{R}) \Rightarrow f$ is finite valued a.e.
3. (Linearity) For $f, g \in L^1(\mathbb{R})$ and $a, b \in \mathbb{R}$, $af + bg \in L^1(\mathbb{R})$ and $\int_{\mathbb{R}}(af + bg) = a\int_{\mathbb{R}} f + b\int_{\mathbb{R}} g$
4. If $f \in L^1(\mathbb{R})$ and $A \in \mathcal{M}$ and $m(A) = 0$ then $\int_A f = 0$; in particular if $f, g \in L^1(\mathbb{R})$ with $f = g$ a.e. then $\int_{\mathbb{R}} f = \int_{\mathbb{R}} g$
5. (Monotonicity) If $f, g \in L^1(\mathbb{R})$ with $f \leq g$ a.e., then $\int_{\mathbb{R}} f \leq \int_{\mathbb{R}} g$

Proof.

1. $-\int_{\mathbb{R}} f^- \leq \int_{\mathbb{R}} f \leq \int_{\mathbb{R}} f^+$ and $\int_{\mathbb{R}} f^{\pm} \leq \int_{\mathbb{R}} |f|$.
2. We know $\int_{\mathbb{R}} |f| < \infty$ so $|f| < \infty$ a.e. by properties of integrals of non-negative functions so $m(\{f = \pm\infty\}) = 0$
3. $|af| \leq |a|\,|f|$ so by monotonicity of non-negative functions, $\int_{\mathbb{R}} |af| \leq |a| \int_{\mathbb{R}} |f| < \infty$ so af in $L^1(\mathbb{R})$. Note then that

$$(af)^+ = \begin{cases} af^+ \text{ if } a \geq 0 \\ -af^- \text{ if } a < 0 \end{cases}, \qquad (af)^- = \begin{cases} af^- \text{ if } a \geq 0 \\ -af^+ \text{ if } a < 0 \end{cases}$$

so

$$\begin{aligned} \int_{\mathbb{R}} af &= \int_{\mathbb{R}} (af)^+ - \int_{\mathbb{R}} (af)^- \\ &= \begin{cases} \int_{\mathbb{R}} af^+ - \int_{\mathbb{R}} af^- \text{ if } a \geq 0 \\ \int_{\mathbb{R}} (-a)f^- - \int_{\mathbb{R}} (-a)f^+ \text{ if } a < 0 \end{cases} \\ &= \begin{cases} a\left(\int_{\mathbb{R}} f^+ - \int_{\mathbb{R}} f^-\right) \text{ if } a \geq 0 \\ (-a)\left(\int_{\mathbb{R}} f^- - \int_{\mathbb{R}} f^+\right) \text{ if } a < 0 \end{cases} = a\int_{\mathbb{R}} f. \end{aligned}$$

By the same argument $bg \in L^1(\mathbb{R})$ and $\int_{\mathbb{R}} (bg) = b\int_{\mathbb{R}} g$. wlog, $a = b = 1$. We want to show $f + g \in L^1(\mathbb{R})$; clearly $|f + g| \leq |f| + |g| < \infty$ so it must be $f + g \in L^1(\mathbb{R})$. Set $h := f + g$ then $|h, f, g| < \infty$ a.e. and each of the integrals of $|h, f, g| < \infty$. Then, $h^+ - h^- = f^+ - f^- + g^+ - g^-$. Then $h^+ + f^- + g^- = f^+ + g^+ + h^-$, where now both sides are non-negative functions. By linearity of integrals of non-negative functions and since all terms finite a.e.,

$$\begin{aligned}\int h^+ + \int f^- + \int g^- = \int f^+ + \int g^+ + \int h^- \\ \Rightarrow \int h^+ - \int h^- = \int f^+ - \int f^- + \int g^+ - \int g^- \\ \Rightarrow \int (f+g) = \int h = \int f + \int g.\end{aligned}$$

4. $|\int_A f| \leq \int_A |f| = 0$.
5. Put $h = g - f$ (valid since $f, g \in L^1(\mathbb{R})$) then $h \geq 0$ a.e. Then $\int_{\mathbb{R}} h \geq 0$ so by linearity $\int_{\mathbb{R}} (g - f) = \int_{\mathbb{R}} g - \int_{\mathbb{R}} f \geq 0$.

■

§2.6 Convergence Theorems of Integral

↪**Theorem 2.7** (Monotone Covergence Theorem (MON)): Assume $\{f_n\}, f$ are non-negative, measurable functions. If $f_n \uparrow$ and $\lim_{n\to\infty} f_n = f$, then

$$\int_{\mathbb{R}} f = \lim_{n\to\infty} \int_{\mathbb{R}} f_n.$$

Remark 2.8: When we write $\lim_n f_n = f$, we mean pointwise convergence; however, one can replace these statements with convergence a.e. and obtain an equivalent, more general result wlog.

PROOF. By monotonicity of non-negative functions, $\lim_{n\to\infty} \int_{\mathbb{R}} f_n$ exists, forming an increasing sequence. Since $f_n \leq f$, then we know too that $\lim_{n\to\infty} \int_{\mathbb{R}} f_n \leq \int_{\mathbb{R}} f$.

Conversely, for every n, let $\{\varphi_{n,k}\}_{k\in\mathbb{N}}$ be a sequence of simple functions such that $\varphi_{n,k} \uparrow$ w.r.t k and $\varphi_{n,k} \to f_n$ as $k \to \infty$;

$$\begin{array}{ccccccc} f_1 & f_2 & \cdots & f_k & f_{k+1} & \cdots & \to f \\ \vdots & \vdots & \iddots & \vdots & \vdots & & \\ \varphi_{1,k} & \varphi_{2,k} & \iddots & \varphi_{k,k} & \varphi_{k+1,k} & \cdots & \\ \vdots & \vdots & \iddots & \vdots & \vdots & \cdots & \\ \varphi_{1,2} & \varphi_{2,2} & \iddots & \varphi_{k,2} & \varphi_{k+1,2} & \cdots & \\ \varphi_{1,1} & \varphi_{2,1} & \cdots & \varphi_{k,1} & \varphi_{k+1,1} & \cdots & \end{array}$$

For each $k \geq 1$, let

$$g_k := \max\{\varphi_{1,k}, \varphi_{2,k}, \ldots, \varphi_{k,k}\}.$$

Then, g_k simple for each k, and $g_k \uparrow$ and $g_k \leq f$. So, $\lim_{k\to\infty} g_k$ exists. Then, for all $n \geq 1$, $\lim_{k\to\infty} g_k \geq \lim_{k\to\infty} \varphi_{n,k} = f_n$ so $\lim_{k\to\infty} g_k \geq \lim_{n\to\infty} f_n = f$. Thus, $\lim_{k\to\infty} \int_{\mathbb{R}} g_k = \int_{\mathbb{R}} f$ by a previous theorem. Since $\forall\, k \geq 1$, $\varphi_{1,k}, \varphi_{2,k}, \cdots, \varphi_{k,k} \leq f_k$, $g_k \leq f_k$ and thus by monotonicity $\int_{\mathbb{R}} g_k \leq \int_{\mathbb{R}} f_k \Rightarrow \int_{\mathbb{R}} f = \lim_{k\to\infty} \int_{\mathbb{R}} g_k \leq \lim_{k\to\infty} \int_{\mathbb{R}} f_k$ as desired. ■

↪**Corollary 2.5**: If $\{f_n\}, f$ measurable functions such that $f_n \uparrow$ and $\lim_n f_n = f$ and $\int_{\mathbb{R}} f_1^- < \infty$, then $\int_{\mathbb{R}} f = \lim_n \int_{\mathbb{R}} f_n$.

Proof. Since $f_n \uparrow, f_n \geq f_1$ so $f \geq f_1$. Then, $f_n^- \leq f_1^-, f^- \leq f_1^-$, all of these are finite valued a.e., and $\int_{\mathbb{R}} f_n^- \leq \int_{\mathbb{R}} f_1^- < \infty$ and $\int_{\mathbb{R}} f^- \leq \int_{\mathbb{R}} f_1^- < \infty$. For each $n \geq 1$, set $\tilde{f}_n := f_n + f_1^- = f_n^+ - f_n^- + f_1^- \geq 0$, and $\tilde{f}_n \uparrow$ with $\lim_n \tilde{f}_n = f + f_1^- =: \tilde{f} \geq 0$. By MON, $\int_{\mathbb{R}} \tilde{f} = \lim_n \int_{\mathbb{R}} \tilde{f}_n$ so $\int_{\mathbb{R}} (f + f_1^-) = \lim_n \int_{\mathbb{R}} (f_n + f_1^-)$.

We have that $\tilde{f}_n = f_n + f_1^- = f_n^+ - f_n^- + f_1^- \Rightarrow \tilde{f}_n + f_n^- = f_n^+ + f_1^-$, which is valid since $f_n^- < \infty$ a.e.. By linearity, then,

$$\begin{aligned}
\int_{\mathbb{R}} \tilde{f}_n + \int_{\mathbb{R}} f_n^- &= \int_{\mathbb{R}} f_n^+ + \int_{\mathbb{R}} f_1^- \\
\Rightarrow \int_{\mathbb{R}} \tilde{f}_n &= \int_{\mathbb{R}} f_n^+ - \int_{\mathbb{R}} f_n^- + \int_{\mathbb{R}} f_1^- \qquad \text{because } \int_{\mathbb{R}} f_n^- < \infty \\
\Rightarrow \int_{\mathbb{R}} \tilde{f}_n &= \int_{\mathbb{R}} f_n + \int_{\mathbb{R}} f_1^-.
\end{aligned}$$

Similar work gives $\int_{\mathbb{R}} \tilde{f} = \int_{\mathbb{R}} f + \int_{\mathbb{R}} f_1^-$, and taking limits and using $\lim_n \int_{\mathbb{R}} (f_n + f_1^-) = \int_{\mathbb{R}} (f + f_1^-)$ completes the proof. ■

↪**Theorem 2.8** (Reverse MON): Assume $\{f_n\}$, measurable such that $f_n \downarrow$ and $\lim_{n\to\infty} f_n = f$. If $\int_{\mathbb{R}} f_1^+ < \infty$, then $\int_{\mathbb{R}} f = \lim_n \int_{\mathbb{R}} f_n$.

Proof. Consider $\{-f_n\}$ and use the previous corollary. ■

↪**Theorem 2.9** (Fatou's Lemma): Assume $\{f_n\}$ non-negative, measurable. Then

$$\int_{\mathbb{R}} \left(\liminf_{n\to\infty} f_n \right) \leq \liminf_{n\to\infty} \left(\int_{\mathbb{R}} f_n \right).$$

Proof. For every $m \geq 1$, set $g_m := \inf_{n \geq m} f_n$. Then, g_m non-negative and $g_m \uparrow$, with $\lim_m g_m = \liminf_n f_n$. By MON, $\int_{\mathbb{R}} \liminf_n f_n = \lim_{m\to\infty} \left(\int_{\mathbb{R}} g_m \right)$. For every $n \geq m$, $g_m \leq f_n$, so by monotonicity, $\int_{\mathbb{R}} g_m \leq \int_{\mathbb{R}} f_n$ for every $n \geq m$, so $\int_{\mathbb{R}} g_m \leq \inf_{n \geq m} \int_{\mathbb{R}} f_n$, and hence $\lim_{m\to\infty} \int_{\mathbb{R}} g_m \leq \lim_{m\to\infty} \inf_{n\geq m} \int_{\mathbb{R}} f_n = \liminf_n \left(\int_{\mathbb{R}} f_n \right)$, and the proof follows. ■

↪**Corollary 2.6**: Assume $\{f_n\}$ measurable and there exists a measurable function g such that $\int_{\mathbb{R}} g^- < \infty$ and $f_n \geq g$ for every n. Then,

$$\int_{\mathbb{R}} \left(\liminf_n f_n \right) \leq \liminf_n \left(\int_{\mathbb{R}} f_n \right).$$

Proof. Since $f_n \geq g$ for all $n \geq 1$, $f_n^- \leq g^-$ so $f_n^- < \infty$ a.e. and $\int_{\mathbb{R}} f_n^- < \infty$. Set $\tilde{f}_n := f_n + g^- \geq 0$. Then, apply Fatou to get $\int_{\mathbb{R}} \liminf_n \tilde{f}_n \leq \liminf_n \int_{\mathbb{R}} \tilde{f}_n$, then it suffices to check linearity. ■

↪**Theorem 2.10** (Reverse Fatou): Assume $\{f_n\}$ measurable and there exists a g measurable such that $\int_{\mathbb{R}} g^+ < \infty$ and $f_n \le g$ for all $n \ge 1$. Then,

$$\int_{\mathbb{R}} \left(\limsup_n f_n \right) \ge \limsup_n \left(\int_{\mathbb{R}} f_n \right).$$

Proof. Apply previous proof to $\{-f_n\}$. ■

Remark 2.9: The "floor" g is necessary. Let $f_n(x) := \begin{cases} -1 & \text{if } x \ge n \\ 0 & \text{if } x < n \end{cases}$. Then, $f_n \uparrow$, and $\lim_n f_n = 0$ while $\int_{\mathbb{R}} f_n = -\infty$ for every n, so MON doesn't apply.

↪**Theorem 2.11** (Dominated Convergence Theorem (DOM)): Assume $\{f_n\}, f$ measurable with $\lim_n f_n = f$. If there exists a $g \in L^1(\mathbb{R})$ such that $|f_n| \le |g|$ for all n, then $f_n \to f$ in $L^1(\mathbb{R})$ i.e. $\lim_{n\to\infty} \int_{\mathbb{R}} |f_n - f| = 0$. In particular, $\int_{\mathbb{R}} f = \lim_n \int_{\mathbb{R}} f_n$.

Proof. Since $|f_n| \le |g|$ and $f = \lim_{n\to\infty} f_n$, then $|f| \le |g|$. So, $\int_{\mathbb{R}} |f_n| \le \int_{\mathbb{R}} |g| < \infty$ and similarly $\int_{\mathbb{R}} |f| \le \int_{\mathbb{R}} |g| < \infty$ so $|f_n|, f \in L^1(\mathbb{R})$.

Observe that $|f_n - f| \le 2\,|g|$, and $\int_{\mathbb{R}} (2\,|g|) < \infty$. Applying Reverse Fatou to $\{|f_n - f|\}_{n\in\mathbb{N}}$, we find

$$\int_{\mathbb{R}} \left(\underbrace{\limsup_n (|f_n - f|)}_{0} \right) \ge \limsup_n \left(\int_{\mathbb{R}} |f_n - f| \right)$$

$$\Rightarrow \lim_{n\to\infty} \int_{\mathbb{R}} |f_n - f| = 0,$$

so in particular

$$\left| \int_{\mathbb{R}} f_n - \int_{\mathbb{R}} f \right| = \left| \int_{\mathbb{R}} (f_n - f) \right| \le \int_{\mathbb{R}} |f_n - f| \to 0$$

so $\lim_n \int_{\mathbb{R}} f_n = \int_{\mathbb{R}} f$. ■

Remark 2.10: We must find $g \in L^1(\mathbb{R})$ to dominate $|g| \ge |f_n|$ irrespective of n. For instance, if $f_n = \mathbb{1}_{[n,2n]}$, then $\lim_n f_n = 0$, but $\int_{\mathbb{R}} f_n = n$ for all $n \ge 1$. DOM doesn't apply, since we would need a constant 1 function to dominate all f_n, which is not integrable.

↪**Proposition 2.22**: Assume $f \in L^1(\mathbb{R})$, $\{h_n\}$ a sequence of measurable functions that are uniformly bounded, i.e. $\exists\, M > 0$ such that $|h_n| \le M$ a.e. for all $n \ge 1$. If $h_n \to h$ a.e. for some measurable function h, then

$$\lim_n \int_{\mathbb{R}} (f h_n) = \int_{\mathbb{R}} (f h).$$

Proof. For every n, $|f \cdot h_n| \leq M\,|f| \in L_1(\mathbb{R})$. The conclusion follows from DOM. ■

↪**Corollary 2.7**: If $f \in L^1(\mathbb{R})$ then for all $\varepsilon > 0$, there exists a compact set $K \subseteq \mathbb{R}$ such that $\int_{K^c} |f| \leq \varepsilon$.

Proof. If $h_n := \mathbb{1}_{[-n,n]}$, the $\lim_n \int_{\mathbb{R}} f h_n = \lim_n \int_{[-n,n]} f = \int_{\mathbb{R}} f$, and also $\lim_n \int_{\{\mathbb{R}-[-n,n]\}} f = 0$. ■

↪**Corollary 2.8**: If $f \in L^1(\mathbb{R})$, then for all $\varepsilon > 0$, $\exists\, N \geq 1$ such that $\int_{\{|f|>N\}} |f| \leq \varepsilon$.

Proof. Let $h_n = \mathbb{1}_{\{|f|>n\}}$ then $\lim_{n\to\infty} \int_{\{|f|>n\}} f = 0$. ■

↪**Corollary 2.9**: If $\{A_n\} \subseteq \mathcal{M}$ such that $A_n \uparrow$, then $\int_{\bigcup_{n=1}^{\infty} A_n} f = \lim_{n\to\infty} \int_{A_n} f$ $(\mathbb{1}_{A_n} f \to \mathbb{1}_{\bigcup_{n=1}^{\infty} A_n} f)$.

↪**Corollary 2.10** (Countable Additivity): If $\{B_n\} \subseteq \mathcal{M}$ are disjoint, then $\int_{\bigcup_{n=1}^{\infty} B_n} f = \sum_{n=1}^{\infty} \int_{B_n} f$.

↪**Corollary 2.11**: If $\{A_n\} \subseteq \mathcal{M}$ such that $A_n \downarrow$, then $\int_{\bigcap_{n=1}^{\infty} A_n} f = \lim_{n\to\infty} \int_{A_n} f$.

↪**Proposition 2.23**: Assume f is non-negative, measurable, and finite-valued a.e.. Then, for every $k \in \mathbb{Z}$, put $A_k := \left\{x \in \mathbb{R} : 2^k \leq f(x) < 2^{k+1}\right\}$. Then,

$$f \text{ integrable} \Leftrightarrow \int_{\mathbb{R}} f < \infty \Leftrightarrow \sum_{k\in\mathbb{Z}} 2^k m(A_k) < \infty.$$

Proof. ($\Rightarrow$) Note that the A_k's disjoint and $\bigcup_{k\in\mathbb{Z}} A_k = \{0 < f < \infty\}$. So,

$$\int_{\mathbb{R}} f = \underbrace{\int_{\{f=0\}} f}_{=0 \text{ since } f=0} + \int_{\{0<f<\infty\}} + \underbrace{\int_{\{f=\infty\}} f}_{=0 \text{ since } f<\infty \text{ a.e.}} = \sum_{k\in\mathbb{Z}} \int_{A_k} f.$$

For each $k \in \mathbb{Z}$, for every $x \in A_k$, $2^k \leq f(x) < 2^{k+1}$ so $2^k m(A_k) \leq \int_{A_k} f(x) < 2^{k+1} m(A_k)$. Hence,

$$\sum_{k\in\mathbb{Z}} 2^k m(A_k) \leq \sum_{k\in\mathbb{Z}} \int_{A_k} f = \int_{\mathbb{R}} f < \infty.$$

($\Leftarrow$) Suppose $\sum_{k\in\mathbb{Z}} 2^k m(A_k) < \infty$. We know again

$$\int_{\mathbb{R}} f = \int_{\{0<f<\infty\}} f \underset{\text{By MON}}{=} \sum_{k\in\mathbb{Z}} \int_{A_k} f < \sum_{k\in\mathbb{Z}} 2^{k+1} m(A_k) = 2 \sum_{k\in\mathbb{Z}} 2^k m(A_k) < \infty.$$

■

⊛ **Example 2.3**: Let $f(x) = |x|^{-\alpha} \mathbb{1}_{[-1,1]}(x)$, with $f(0) = \infty$ and $\alpha > 0$; f finite-valued a.e.. For every $k \in \mathbb{Z}$, put $A_k := \left\{2^k \leq f < 2^{k+1}\right\} = \left\{x \in [-1,1] : 2^k \leq |x|^{-\alpha} < 2^{k+1}\right\}$. By definition, $|f| \geq 1$, so

$$A_k = \left[-2^{-\frac{k}{\alpha}}, -2^{\frac{-(k+1)}{\alpha}}\right) \cup \left(2^{\frac{-(k+1)}{\alpha}}, 2^{-\frac{k}{\alpha}}\right] \text{ for } k \geq 0, \qquad A_k = \varnothing \text{ if } k < 0.$$

Hence,

$$\sum_{k \in \mathbb{Z}} 2^k m(A_k) = \sum_{k=0}^{\infty} 2^k \cdot 2 \cdot \left(1 - 2^{-\frac{1}{\alpha}}\right) 2^{-\frac{k}{\alpha}} = 2\left(1 - 2^{-\frac{1}{\alpha}}\right) \sum_{k=0}^{\infty} 2^{k\left(1 - \frac{1}{\alpha}\right)}.$$

Hence, the series $< \infty \Leftrightarrow \alpha < 1$, and thus $\int_{[-1,1]} |x|^{-\alpha} \, \mathrm{d}x < \infty \Leftrightarrow \alpha < 1$.

⊛ **Example 2.4**: Let $g(x) = |x|^{-\beta} \mathbb{1}_{\mathbb{R}-[-1,1]}(x)$ with $\beta > 0$. We have $|g| < 1$; we again put

$$A_k := \left\{2^k \leq g < 2^{k+1}\right\} = \begin{cases} \left[-2^{-\frac{k}{\beta}}, -2^{\frac{-(k+1)}{\beta}}\right) \cup \left(2^{\frac{-(k+1)}{\beta}}, 2^{-\frac{k}{\beta}}\right] & \text{if } k < 0 \\ \varnothing & \text{if } k \geq 0. \end{cases}.$$

So,

$$\int_{\mathbb{R}-[-1,1]} |x|^{-\beta} \, \mathrm{d}x < \infty \Leftrightarrow \sum_{k \in \mathbb{Z}} 2^k m(A_k) < \infty \Leftrightarrow \beta > 1.$$

⊛ **Example 2.5**: Let $f_n(x) = \left(1 + \frac{x}{n}\right)^{-n} \sin\left(\frac{x}{n}\right)$. What is $\lim_{n \to \infty} \int_{(0,\infty)} f_n(x) \, \mathrm{d}x$? We have that for all $x > 0$, $\lim_{n \to \infty} f_n(x) = 0$. We have that since $|\sin\left(\frac{x}{n}\right)| \leq 1$, so

$$|f_n(x)| \leq \left(1 + \frac{x}{n}\right)^{-n} \leq \left(1 + \frac{x}{2}\right)^{-2} \; \forall\, x > 0, \forall\, n \geq 2.$$

Let $g(x) := \left(1 + \frac{x}{2}\right)^{-2}$. We would like to apply DOM, so we need to check that $g \in L^1((0,\infty))$. We have that

$$\int_{(0,\infty)} g = \int_{(0,1]} g + \int_{(1,\infty)} g \leq \int_{(0,1]} 1 + \underbrace{\int_{(1,\infty)} \frac{4}{x^2} \, \mathrm{d}x}_{\beta=2 \text{ of previous example}} < \infty,$$

so indeed $g \in L^1((0,\infty))$. Applying DOM, then, we have that

$$\lim_{n \to \infty} \int_{(0,\infty)} f_n = \int_{(0,\infty)} \lim_{n \to \infty} f_n = 0.$$

⊛ **Example 2.6**: Let $c > 0, f_n(x) = x^{-c}(\cosh x)^{-\frac{1}{n}}$. What is $\lim_n \int_{(1,\infty)} f_n$?

For every $x > 1$, $\cosh x > 1$, so $(\cosh x)^{-\frac{1}{n}} \uparrow$ with respect to n, with $\lim_n (\cosh x)^{-\frac{1}{n}} = 1$, so $\lim_{n\to\infty} f_n(x) = x^{-c}$ for every $x > 1$. Let $g(x) = x^{-c}$, then. By previous examples, when $c > 1$, $g \in L^1((1,\infty))$ so DOM applies and thus

$$\lim_n \int_{(1,\infty)} f_n = \int_{(1,\infty)} \lim_n f_n = \int_{(1,\infty)} x^{-c}\,\mathrm{d}x < \infty.$$

When $0 < c \leq 1$, by Fatou,

$$\liminf_n \int_{(1,\infty)} f_n \geq \int_{(1,\infty)} \liminf_n (f_n) = \int_{(1,\infty)} x^{-c}\,\mathrm{d}x,$$

since f_n converges. When $0 < c \leq 1$, the RHS $= \infty$, and thus $\lim_{n\to\infty} \int_{(1,\infty)} f_n = \infty$.

⊛ **Example 2.7**: Let $c \geq 0, f_n(x) := \frac{n}{1+n^2x^2}$ for $x \geq 0$. What is $\lim_n \int_{[c,\infty)} f_n$?

We have that

$$\lim_n f_n(x) = \begin{cases} 0 & \text{if } x > 0 \\ \infty & \text{if } x = 0 \end{cases}.$$

On $x \in [1,\infty), f_n(x) \geq f_{n+1}(x)$ for all $n \geq 1$, namely $f_n \downarrow$, and so $f_n(x) \leq f_1(x) = \frac{1}{1+x^2}$. $f_1(x) \in L^1(\mathbb{R})$, by comparison with $\frac{1}{x^2}$ $(\alpha = 2)$.

If $x \in (0,1), f_n(x) = \frac{1}{x}\frac{nx}{1+(nx)^2} \leq A\frac{1}{x}$, with $A := \sup_{t>0} \frac{t}{1+t^2} < \infty$. But $\frac{A}{x} \notin L^1((0,1))$.

When $c > 0$, for all $x \geq c$ and for all $n \geq 1$,

$$f_n(x) \leq \mathbb{1}_{[1,\infty)}(x)\frac{1}{1+x^2} + \mathbb{1}_{[c,1)}\frac{A}{x} \leq \mathbb{1}_{[1,\infty)}(x)\frac{1}{1+x^2} + \mathbb{1}_{[c,1)}(x)\frac{A}{c} \in L^1([c,\infty)).$$

Hence, we may apply DOM, so

$$\lim_n \int_{[c,\infty)} f_n = \int_{[c,\infty)} \lim_n f_n = 0,$$

when $c > 0$. However, when $c = 0$, we have no such dominating function; so what is $\int_{[0,\infty)} f_n(x)\,\mathrm{d}x$?

§2.7 Riemann Integral vs Lebesgue Integral

Recall; let f be bounded on $[a,b]$. Then, f is Riemann integrable on $[a,b]$ if

$$\begin{cases} f \text{ is continuous on } [a,b] \\ f \text{ is monotonic on } [a,b] \\ f \text{ is continuous except at possibly finitely many points in } [a,b] \end{cases}.$$

Recall the function $f = \mathbb{1}_{\mathbb{Q}\cap[0,1]}$. f is not Riemann integrable, but is Lebesgue integrable, because $|f| \leq \mathbb{1}_{[0,1]} \in L^1(\mathbb{R})$.

Remark 2.11:
1. $\exists$ bounded functions on $[a,b]$ that are not Riemann integrable.
2. In general, g being Riemann integrable and $|f| \leq |g| \not\Rightarrow f$ is Riemann integrable ($\mathbb{1}_{\mathbb{Q}\cap[0,1]} \leq \mathbb{1}_{[0,1]}$).
3. In general, DOM and MON do *not* apply to Riemann integrable. For instance, consider $\{q_n\}$ an enumeration of $\mathbb{Q} \cap [0,1]$. Define $f_n(x) := \begin{cases} 1 \text{ if } x \in \{q_1,...,q_n\} \\ 0 \text{ else} \end{cases}$. $f_n \uparrow$, with $f_n \to \mathbb{1}_{\mathbb{Q}\cap[0,1]}$. So, MON applies with the Lebesgue integral, but f_n is only discontinuous, for every n, at finitely many points, so f_n Riemann integrable with $\int_0^{1^{(R)}} f_n = 0$, but the limit is not Riemann integrable.

↪**Theorem 2.12**: Assume f is Riemann integrable on $[a,b]$. Then, f is Lebesgue integrable on $[a,b]$, i.e. $f \in L^1([a,b])$. Moreover, $\int_a^{b^{(R)}} f = \int_{[a,b]} f$.

Proof. f is Riemann integrable on $[a,b]$, so there is some $M > 0$ such that $|f| \leq M$ on $[a,b]$. Further, there exist step functions φ_n, ψ_n with $\varphi_n \leq f \leq \psi_n$ on $[a,b]$ and $|\varphi_n|, |\psi_n| \leq M$ for all $n \geq 1$, and

$$\lim_{n\to\infty} \int_a^{b^{(R)}} \varphi_n = \int_a^{b^{(R)}} f = \lim_{n\to\infty} \int_a^{b^{(R)}} \psi_n.$$

Denote $\varphi := \lim_{n\to\infty} \varphi_n, \psi := \lim_{n\to\infty} \psi_n$, which exist by Monotonicity. Since φ_n, ψ_n are step functions, they are measurable hence φ, ψ measurable with $\varphi \leq f \leq \psi$. Observe that the Lebesgue, Riemann integral coincide on step functions. Hence, $\int_a^{b^{(R)}} \varphi_n = \int_{[a,b]} \varphi_n$, same with ψ_n. By DOM, (with M as the dominator)

$$\int_{[a,b]} \varphi = \lim_n \int_{[a,b]} \varphi_n = \lim_n \int_a^{b^{(R)}} \varphi_n = \int_a^{b^{(R)}} (f) = \lim_n \int_a^{b^{(R)}} \psi_n = \lim_n \int_{[a,b]} \psi_n = \int_{[a,b]} \psi.$$

Since $\varphi \leq \psi$ and $\int_{[a,b]}(\psi - \varphi) = 0$, we have that $\psi = \varphi$ a.e. on $[a,b]$ by properties of integrals of non-negative functions, and thus $f = \varphi = \psi$ a.e. on $[a,b]$. In particular, then, f is measurable, being equal a.e. to measurable functions. Thus, since $|f| \leq M$ on $[a,b]$, $f \in L^1([a,b])$, and so since integrals agree on functions that are equal a.e., $\int_{[a,b]} f = \int_{[a,b]} \varphi = \int_a^{b^{(R)}} f$ as desired. ∎

⊛ **Example 2.8**: We return to our example of computing $\lim_{n\to\infty} \int_{[0,\infty)} \frac{n}{1+n^2x^2}\,\mathrm{d}x$. We may rewrite

$$\int_{[0,\infty)} \frac{n}{1+n^2x^2}\,\mathrm{d}x = \int_{[0,T]} \frac{n}{1+n^2x^2}\,\mathrm{d}x + \int_{[T,\infty)} \frac{n}{1+n^2x^2}\,\mathrm{d}x$$

where $T > 0$. We know from the previous example that the RHS integral converges to 0 by application of DOM. Now, $\frac{n}{1+n^2x^2}$ is continuous on $[0,T]$ and thus Riemann integrable, and so by the previous theorem

$$\int_{[0,T]} \frac{n}{1+n^2x^2} = \int_{[0,T]}^{(R)} \frac{n}{1+n^2x^2} = \arctan(nT).$$

As $n \to \infty$, $\arctan(nT) \to \frac{\pi}{2}$, and thus the limit of the whole integral indeed exists, and is in fact equal to $\frac{\pi}{2}$.

§2.8 L^p-space

↪**Definition 2.12** (p-integrable): Let f measurable and $1 \leq p < \infty$. We say f is *p-integrable* and write $f \in L^p(\mathbb{R})$ if $\int_{\mathbb{R}} |f|^p < \infty$, i.e. $|f|^p \in L^1(\mathbb{R})$.

For $f \in L^p(\mathbb{R})$, define the p-norm

$$\|f\|_p := \left(\int_{\mathbb{R}} |f|^p\right)^{\frac{1}{p}}.$$

Remark 2.12: When $p = 1$, we see that $\|\cdot\|_1$ a norm fairly clearly from properties of the integral. We need to show this for more general $p > 1$.

Remark 2.13: $\|\cdot\|_p$ also defined when $p = \infty$; given f measurable, we define

$$\|f\|_\infty := \operatorname{ess\,sup}_{x\in\mathbb{R}} |f(x)| := \inf\left\{a \in \overline{\mathbb{R}} : |f| \leq a \text{ a.e.}\right\}.$$

Then, we define

$$L^\infty(\mathbb{R}) := \{f \text{ measurable s.t. } \|f\|_\infty < \infty\}.$$

One can show that if $f \in L^\infty(\mathbb{R})$, $|f| \leq \|f\|_\infty$ a.e..

↪**Theorem 2.13** (Hölder's Inequality): Let $1 < p < \infty$ and let $q := \frac{p}{p-1}$ (such a q is called the *Hölder Conjugate* of p). If $f \in L^p(\mathbb{R})$ and $g \in L^q(\mathbb{R})$, then $fg \in L^1(\mathbb{R})$, and

$$\|fg\|_1 \leq \|f\|_p \, \|g\|_q.$$

In particular, if $p = q = 2$, then we have the *Cauchy-Schwarz Inequality*.

Remark 2.14: $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. We will employ "Young's Inequality", which states that for all $a, b \geq 0$, $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$ where $\frac{1}{p} + \frac{1}{q} = 1$. Since $f \in L^p, g \in L^q$, set $\tilde{f} := \frac{f}{\|f\|_p}$ and $\tilde{g} := \frac{g}{\|g\|_q}$. Then, a.e.

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

We have

$$\int_{\mathbb{R}} |\tilde{f}\tilde{g}| = \int_{\mathbb{R}} \frac{|fg|}{\|f\|_p \|g\|_q}$$

and

$$\int_{\mathbb{R}} \frac{|\tilde{f}|^p}{p} + \frac{|\tilde{g}|^q}{q} = \frac{1}{p} \frac{\int_{\mathbb{R}} |f|^p}{\|f\|_p^p} + \frac{1}{q} \frac{\int_{\mathbb{R}} |g|^q}{\|g\|_q^q} = \frac{1}{p} + \frac{1}{q} = 1,$$

and thus

$$\int_{\mathbb{R}} |fg| = \|fg\|_q \leq \|f\|_p \|g\|_q$$

as required. ■

Remark 2.15: This inequality also holds for $p = 1, q = \infty$ (assignment question).

↪**Lemma 2.2**: For all $a, b \geq 0$, $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$ where $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. ■

↪**Theorem 2.14** (Minkowski's Inequality): Let $1 \leq p < \infty$ and $f, g \in L^p(\mathbb{R})$. Then, $f + g \in L^p(\mathbb{R})$, and in particular

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

In particular, then, $\|\cdot\|_p$ satisfies the triangle inequality and is indeed a norm on $L^p(\mathbb{R})$.

Proof. We have $|f + g|^p \leq 2^p(|f|^p + |g|^p)$ hence $f + g \in L^p(\mathbb{R})$ since $|f|^p, |g|^p \in L^1(\mathbb{R})$. Further

$$
\begin{aligned}
\int_{\mathbb{R}} |f+g|^p = \int_{\mathbb{R}} |f+g|\,|f+g|^{p-1} &\le \int_{\mathbb{R}} |f|\,|f+g|^{p-1} + \int_{\mathbb{R}} |g|\,|f+g|^{p-1} \\
\text{(Hölder's)} \qquad &\le \left(\int_{\mathbb{R}} |f|^p\right)^{\frac{1}{p}} \left(\int_{\mathbb{R}} |f+g|^{(p-1)q}\right)^{\frac{1}{q}} + \left(\int_{\mathbb{R}} |g|^p\right)^{\frac{1}{p}} \left(\int_{\mathbb{R}} |f+g|^{(p-1)q}\right)^{\frac{1}{q}} \\
&\le \left(\|f\|_p + \|g\|_p\right)\left(\int_{\mathbb{R}} |f+g|^p\right)^{\frac{1}{q}} \\
&\Rightarrow \|f+g\|_p = \left(\int_{\mathbb{R}} |f+g|^p\right)^{\frac{1}{p}} = \left(\int_{\mathbb{R}} |f+g|^p\right) \cdot \left(\int_{\mathbb{R}} |f+g|^p\right)^{-\frac{1}{q}} \\
&\qquad \le \left(\|f\|_p + \|g\|_p\right)\left(\int_{\mathbb{R}} |f+g|^p\right)^{\frac{1}{q}} \cdot \left(\int_{\mathbb{R}} |f+g|^p\right)^{-\frac{1}{q}} = \|f\|_p + \|g\|_p \\
&\Rightarrow \|f+g\|_p \le \|f\|_p + \|g\|_p
\end{aligned}
$$

■

Remark 2.16: Minkowski's also holds for $p = \infty$.

↪**Lemma 2.3**: Let $1 \le p < \infty$. If $\{g_k\} \in L^p(\mathbb{R})$ such that $\sum_{k=1}^{\infty} \|g_k\|_p < \infty$, then $\exists\, G \in L^p(\mathbb{R})$ such that $G_m := \sum_{k=1}^{m} g_k \to G$ as $m \to \infty$ a.e. as well as in $L^p(\mathbb{R})$.

Proof. Put $\widetilde{G_m} := \sum_{k=1}^{m} |g_k|$ and $\tilde{G} := \sum_{k=1}^{\infty} |g_k|$. Then, $\widetilde{G_m} \uparrow$ with $\lim_{m\to\infty} \widetilde{G_m} = \tilde{G}$. By MON,

$$
\int_{\mathbb{R}} \tilde{G}^p = \lim_{m\to\infty} \int_{\mathbb{R}} \widetilde{G_m}^p = \lim_{m\to\infty} \|\widetilde{G_m}\|_p^p \le \lim_{m\to\infty} \left(\sum_{k=1}^{m} \|g_k\|_p\right)^p
$$

where the final inequality is by Minkowski's. Then,

$$
\le \left(\lim_{m\to\infty} \sum_{k=1}^{m} \|g_k\|_p\right)^p = \left(\sum_{k=1}^{\infty} \|g_k\|_p\right)^p < \infty, \text{by assumption}
$$

Hence, $\tilde{G} \in L^p(\mathbb{R})$ and $\|\tilde{G}\|_p \le \sum_{k=1}^{\infty} \|g_k\|_p$ and thus $\tilde{G}$ finite-valued a.e. and hence $\sum_{k=1}^{\infty} g_k$ absolutely convergent a.e.. Set $G = \lim_{m\to\infty} G_m = \sum_{k=1}^{\infty} g_k$ a.e.. Moreover, we know

$$
|G| = \left|\sum_{k=1}^{\infty} g_k\right| \le \sum_{k=1}^{\infty} |g_k| = \tilde{G} \Rightarrow G \in L^p(\mathbb{R})
$$

and

$$
|G - G_m| \le \sum_{k=m+1}^{\infty} |g_k|.
$$

Fix $\varepsilon > 0$. Since $\sum_{k=1}^{\infty} \|g_k\|_p < \infty$, exists some $M \ge 1$ such that $\sum_{k=M+1}^{\infty} \|g_k\|_p < \varepsilon$. Then,

$$\int_{\mathbb{R}} |G - G_M|^p \leq \int_{\mathbb{R}} \left(\sum_{k=M+1}^{\infty} |g_k| \right)^p = \lim_{L \to \infty} \int_{\mathbb{R}} \left(\sum_{k=M+1}^{L} |g_k| \right)^p$$

$$\text{(Minkowski)} \qquad \leq \lim_{L \to \infty} \left(\sum_{k=M+1}^{L} \|g_k\|_p \right)^p$$

$$= \left(\sum_{k=M+1}^{\infty} \|g_k\|_p \right)^p \leq \varepsilon$$

hence $G_m \to G$ in $L^p(\mathbb{R})$. ■

↪**Theorem 2.15**: Let $1 \leq p < \infty$. Then $L^p(\mathbb{R})$ is a complete normed space under the p-norm.

Proof. Let $f_n \in L^p(\mathbb{R})$ be a Cauchy sequence under $\| \cdot \|_p$. We can choose a subsequence $\{n_k\}$ such that for every $k \geq 1$, $\|f_{n_{k+1}} - f_{n_k}\|_p \leq 2^{-k}$. Set $g_k := f_{n_{k+1}} - f_{n_k}$. By the lemma, if $G_m := \sum_{k=1}^{m} g_k$, there exists some $G \in L^p(\mathbb{R})$ such that $G_m \to G$ a.e. and in $L^p(\mathbb{R})$. In fact, we have

$$G_m = \sum_{k=1}^{m} g_k = \sum_{k=1}^{m} \left(f_{n_{k+1}} - f_{n_k} \right) = f_{n_{m+1}} - f_{n_1},$$

hence

$$G = \lim_{m \to \infty} G_m = \left(\lim_{m \to \infty} f_{n_{m+1}} \right) - f_{n_1}.$$

Let $f := G + f_{n_1}$. Then, $f = \lim_{m \to \infty} f_{n_m}$ a.e. and since $G_m \to G$ in L^p, we have that $f_{n_m} \to f$ in L^p as $m \to \infty$. It remains to show convergence in L^p along the whole subsequence.

Fix $\varepsilon > 0$. Let $N \geq 1$ such that $\sup_{k,\ell \geq N} \|f_k - f_\ell\|_p < \varepsilon$ and m sufficiently large such that $n_m > N$ and $\|f_{n_m} - f\|_p \leq \varepsilon$. Then,

$$\|f_n - f\|_p \leq \underbrace{\|f_n - f_{n_m}\|_p}_{< \varepsilon} + \underbrace{\|f_{n_m} - f\|_p}_{< \varepsilon} < 2\varepsilon,$$

completing the proof. ■

Remark 2.17: L^∞ also complete.

2.8.1 Dense Subspaces of $L^p(\mathbb{R})$

↪**Lemma 2.4**: Bounded and compactly supported functions are dense in $L^p(\mathbb{R})$.

Proof. Given $f \in L^p(\mathbb{R})$, set

$$f_n(x) = \mathbb{1}_{[-n,n]}(x) \cdot f(x) \cdot \mathbb{1}_{\{|f| \leq n\}}(x)$$

which are bounded and compactly supported on $[-n, n]$. We claim $f_n \to f$ in $L^p(\mathbb{R})$. We have $\int_{\mathbb{R}} |f_n - f|^p$ nonzero only if $x \notin [-n, n]$ or $|f(x) > n|$. Hence

$$\int_{\mathbb{R}} |f_n - f|^p \leq \int_{\mathbb{R}\setminus[-n,n]} |f|^p + \int_{\{|f|>n\}} |f|^p \to 0 \text{ as } n \to \infty.$$

■

↪**Lemma 2.5**: Simple functions are dense in $L^p(\mathbb{R})$.

PROOF. For $f \in L^p(\mathbb{R})$, let f_n be as in the previous proof. For each $n \geq 1, k = 0, 1, \ldots, n2^n - 1$, set

$$A_{n,k} := \left\{x \in [-n,n] : \frac{k}{2^n} \leq f_n^+ < \frac{k+1}{2^n}\right\}, \qquad \varphi_n^+ := \sum_{k=0}^{n2^n-1} \frac{k}{2^n} \mathbb{1}_{A_{n,k}},$$

and

$$B_{n,k} := \left\{x \in [-n,n] : \frac{k}{2^n} \leq f_n^- < \frac{k+1}{2^n}\right\}, \qquad \varphi_n^- := \sum_{k=0}^{n2^n-1} \frac{k}{2^n} \mathbb{1}_{B_{n,k}}.$$

Put $\varphi_n := \varphi_n^+ - \varphi_n^-$. This is a simple function, and $|\varphi_n| \leq n$ and supported on $[-n,n]$ for every n hence $\varphi_n \in L^p(\mathbb{R})$. In addition, $\lim_n \varphi_n(x) = f(x)$. In particular, for any $n \geq 1$,

$$|f_n(x) - \varphi_n(x)| \leq |f_n^+(x) - \varphi_n^+(x)| + |f_n^-(x) - \varphi_n^-(x)| \leq 2 \cdot 2^{-n}.$$

Then, in particular

$$\|f - \varphi_n\|_p \leq \underbrace{\|f - f_n\|_p}_{\to 0} + \underbrace{\|f_n - \varphi_n\|_p}_{\substack{=\left(\int_{[-n,n]} |f_n - \varphi_n|^p\right)^{\frac{1}{p}} \\ \leq \left((2 \cdot 2^{-n})^p m([-n,n])\right)^{\frac{1}{p}} \to 0}},$$

and so indeed $\varphi_n \to f$ in $L^p(\mathbb{R})$. ■

↪**Theorem 2.16**: Let $C_c(\mathbb{R})$ denote the space of continuous and compactly supported functions. Then, $C_c(\mathbb{R})$ is dense in $L^p(\mathbb{R})$ for $1 \leq p < \infty$.

PROOF. Give $f \in L^p(\mathbb{R})$, let $\{\varphi_n\}$ simple functions as in the previous proof. Recall that, for every $n \geq 1$, there exists a step function θ_n such that $\theta_n \leq \sup_x |\varphi_n(x)| \leq n$, is supported on $[-n-1, n+1]$, *and* $\{\theta_n \neq \varphi_n\}$ has arbitrarily small measure. In particular, we choose θ_n such that $m(\{\theta_n \neq \varphi_n\}) \leq 2^{-n-1}$ for every $n \geq 1$.

Recall that given a step function θ_n, there exists a function $\widetilde{\theta_n}$ continuous on $\mathbb{R}$, $\widetilde{\theta_n}$ is supported on $[-n-2, n+2]$, *and* $m\left(\left\{\widetilde{\theta_n} - \theta_n\right\}\right) \leq 2^{-n-1}$. Thus, $\left\{\widetilde{\theta_n}\right\} \subseteq C_c(\mathbb{R})$, and

$$m\left(\left\{\widetilde{\theta_n} - \varphi_n\right\}\right) \leq m\left(\left\{\widetilde{\theta_n} \neq \theta_n\right\}\right) + m(\{\theta_n \neq \varphi_n\}) \leq 2^{-n}.$$

So, we have

$$\|f - \widetilde{\theta_n}\|_p \leq \underbrace{\|f - \varphi_n\|_p}_{\to 0 \text{ by lemma}} + \underbrace{\|\varphi_n - \widetilde{\theta_n}\|_p}_{\substack{=\left(\int_{\mathbb{R}} |\varphi_n - \widetilde{\theta_n}|^p\right)^{\frac{1}{p}} \\ =\left(\int_{\{\widetilde{\theta_n} \neq \varphi_n\}} |\varphi_n - \widetilde{\theta_n}|^p\right)^{\frac{1}{p}} \\ \leq ((2n)^p 2^{-n})^{\frac{1}{p}} \to 0}},$$

and thus $\widetilde{\theta_n} \to f$ in $L^p(\mathbb{R})$. ■

Remark 2.18: The density of $C_c(\mathbb{R})$ in $L^p(\mathbb{R})$ is useful in the study of properties of generic L^p functions. For instance, show that if $f \in L^p(\mathbb{R})$, then $\lim_{n\to\infty} \int_{\mathbb{R}} |f\left(x + \frac{1}{n}\right) - f(x)|^p \, dx = 0$, that is $f\left(\cdot + \frac{1}{n}\right) \to f$ in $L^p(\mathbb{R})$ using this density.

Remark 2.19: $C_c(\mathbb{R})$ is *NOT* dense in $L^\infty(\mathbb{R})$.

§2.9 Convergence Modes and Uniform Integrability

Recall that, given $\{f_n\}, f$ measurable and finite-valued a.e., we have the following notions of convergence

1. $f_n \to f$ in measure $\Rightarrow \exists \{n_k\}$ such that $f_{n_k} \to f$ a.e. as $k \to \infty$
2. $f_n \to f$ a.e. on $A \in \mathcal{M}$ with $m(A) < \infty \Rightarrow f_n \to f$ in measure on A
3. $f_n \to f$ in $L^p(\mathbb{R})$.

↪**Proposition 2.24**: If $\{f_n\}, f$ in $L^p(\mathbb{R})$ for $1 \leq p < \infty$ and $f_n \to f$ in $L^p(\mathbb{R})$, then $f_n \to f$ in measure.

Proof. For $\delta > 0$, we have

$$m(\{|f_n - f| > \delta\}) = \int_{\{|f_n - f| > \delta\}} 1 \, dx.$$

Remark that $1 \leq \frac{|f_n - f|}{\delta}$ over $\{|f_n - f| > \delta\}$; further $1^p = 1 \leq \left(\frac{|f_n - f|}{\delta}\right)^p$. Hence,

$$\leq \int_{\{|f_n - f| > \delta\}} \frac{|f_n - f|^p}{\delta^p} \, dx \leq \frac{1}{\delta^p} \int_{\mathbb{R}} |f_n - f|^p \leq \frac{1}{\delta^p} \|f_n - f\|_p^p.$$

But by assumption $\|f_n - f\|_p^p \to 0$ for any $\delta > 0$, hence $m(\{|f_n - f| > \delta\}) \to 0$ i.e. $f_n \to f$ in measure. ■

Remark 2.20: In general, convergence in $L^p \not\Rightarrow$ convergence a.e., with the same counter example from convergence in measure $\not\Rightarrow$ convergence a.e..

Remark 2.21: When do we have convergence a.e. $\Rightarrow$ convergence in L^p? This doesn't hold in general, unless some integral convergence theorem from before holds.

Remark 2.22: When do we have convergence in measure ⇒ convergence in L^p? No in general, unless one of the integral convergence theorem holds; with some slight adaptation.

↪**Proposition 2.25** (MON, Measure Version (mMON)): Let f_n non-negative with $f_n \uparrow$ and $f_n \to f$ in measure. Then,

$$\int_{\mathbb{R}} f = \lim_n \int_{\mathbb{R}} f_n.$$

Proof. $f_n \to f$ in measure implies $f_{n_k} \to f$ almost everywhere along some subsequence n_k, so it must be that f non-negative. Suppose the claim fails. Then, there exists some subsequence $\{n_\ell\}$ such that $\int_{\mathbb{R}} f_{n_\ell} \nrightarrow \int_{\mathbb{R}} f$. However, along this subsequence we also have $f_{n_\ell} \to f$ in measure, and hence exists a subsubsequence n_{ℓ_p} such that $f_{n_{\ell_p}} \to f$ a.e.. Then, by MON applied to this subsubsequence, we know that

$$\lim_p \int_{\mathbb{R}} f_{n_{\ell_p}} = \int_{\mathbb{R}} f,$$

a contradiction. ■

↪**Proposition 2.26** (mDOM): If $f_n \in L^1(\mathbb{R})$ with $f_n \to f$ in measure and there exists some $g \in L^1(\mathbb{R})$ such that $|f_n| \leq |g|$, then $f_n \to f$ in $L^1(\mathbb{R})$.

Recall that if $f \in L^1(\mathbb{R})$, then $\int_{\{|f|>n\}} |f| \to$ as $n \to \infty$. The converse does not hold in general; consider $f \equiv 1$. However, we can achieve a partial converse.

For $A \in \mathcal{M}$, we say $f \in L^1(A)$ if $\int_A |f| < \infty$.

↪**Proposition 2.27**: Given $A \in \mathcal{M}$ with $m(A) < \infty$, then

$$f \in L^1(A) \Leftrightarrow \lim_n \int_{A \cap \{|f|>n\}} |f| = 0.$$

Proof. (⇒) We've proven before, c.f. properties of integral of non-negative functions.

(⇐) Choose N such that $\int_{A \cap \{|f|>N\}} |f| \leq 1$. Then,

$$\begin{aligned}\int_A |f| &= \int_{A \cap \{|f| \leq N\}} |f| + \int_{A \cap \{|f|>N\}} |f| \\ &\leq N \cdot m(A) + 1 < \infty.\end{aligned}$$

■

↪**Definition 2.13** (Uniform Integrability): Given $\{f_n\}$ measurable and $A \in \mathcal{M}$, we say $\{f_n\}$ is uniformly integrable on A if

$$\lim_{M\to\infty}\left(\sup_{n\geq 1}\left(\int_{A\cap\{|f_n|>M\}}|f_n|\right)\right) = 0.$$

↪**Proposition 2.28**: Let $\{f_n\}$ measurable, $A \in \mathcal{M}$.

1. If $m(A) < \infty$ and $\{f_n\}$ uniformly integrable on A, then $\{f_n\}$ is bounded in $L^1(A)$, that is $\sup_{n\geq 1}\int_A |f_n| < \infty$.
2. If $\{f_n\}$ is bounded in $L^p(A)$ for any $1 < p < \infty$, then $\{f_n\}$ is uniformly integrable on A.

Proof.

1. Let M such that $\sup_{n\geq 1}\int_{A\cap\{|f_n|>M\}}|f_n| \leq 1$. Then, we have that

$$\begin{aligned}\sup_{n\geq 1}\int_A |f_n| &= \sup_{n\geq 1}\left(\int_{A\cap\{|f_n|\leq M\}}|f_n| + \int_{A\cap\{|f_n|>M\}}|f_n|\right)\\ &\leq M\cdot m(A) + 1 < \infty.\end{aligned}$$

2. For any $M > 0$, note that $1 \leq \left(\frac{|f_n|}{M}\right)^{p-1}$ over $A\cap\{|f_n|>M\}$. So,

$$\begin{aligned}\sup_n \int_{A\cap\{|f_n|>M\}}|f_n| &\leq \sup_n \int_{A\cap\{|f_n|>M\}}|f_n|\left(\frac{|f_n|}{M}\right)^{p-1}\\ &\leq \underbrace{\frac{1}{M^{p-1}}}_{>0}\underbrace{\sup_n\int_A |f_n|^p}_{<\infty} \to 0 \text{ as } M\to\infty.\end{aligned}$$

■

Remark 2.23: Notice that 2. does *not* require finiteness of the measure of A, in particular one can take $A = \mathbb{R}$.

↪**Proposition 2.29**: Given $\{f_n\}$ measurable and $A \in \mathcal{M}$ with $m(A) < \infty$, TFAE:
(i) $f_n \in L^1(A)\ \forall\, n \geq 1, f \in L^1(A)$ and $f_n \to f$ in $L^1(A)$,
(ii) $\{f_n\}$ is uniformly integrable on A and $f_n \to f$ in measure on A.

Proof. (i) ⇒ (ii) Assume $f_n \to f$ in $L^1(A)$, hence $\int_A |f_n| \to \int_A |f|$ so $\{f_n\}$ bounded in $L^1(A)$. Note we've already proven that $f_n \to f$ in measure. For $M > 0$,

$$\int_{A\cap\{|f_n|>M\}} |f_n| \leq \int_{A\cap\{|f_n|>M\}} |f_n - f| + \int_{A\cap\{|f_n|>M\}} |f|$$
$$\leq \underbrace{\int_A |f_n - f|}_{\to 0} + \underbrace{\int_{A\cap\{|f_n|>M\}\cap\left\{|f|\leq\sqrt{M}\right\}} |f|}_{\substack{\leq\sqrt{M}\cdot m(A\cap\{|f_n|>M\}) \\ \leq\sqrt{M}\frac{\sup_n \int_A |f_n|}{M}\to 0 \text{ as } M\to\infty \\ \text{(Markov's)}}} + \underbrace{\int_{A\cap\{|f_n|>M\}\cap\left\{|f|>\sqrt{M}\right\}} |f|}_{\leq\int_{A\cap\left\{|f|>\sqrt{M}\right\}}|f|\to 0 \text{ since } f\in L^1}.$$

Fix $\varepsilon > 0$. Choose N such that for all $n \geq N$, $\int_A |f_n - f| \leq \frac{\varepsilon}{3}$, choose M such that $\int_{A\cap\left\{|f|>\sqrt{M}\right\}} |f| < \frac{\varepsilon}{3}$ and $\frac{\sup_n \int_A |f_n|}{\sqrt{M}} < \frac{\varepsilon}{3}$. Thus,

$$\sup_{n\geq N} \int_{A\cap\{|f_n|>M\}} |f_n| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

We want this to hold for $N = 1$ for uniformity, i.e. we need to deal with the first $N - 1$ terms. We achieve this by making M larger if necessary such that

$$\int_{A\cap\{|f_k|>M\}} |f_k| \leq \varepsilon$$

for every $k = 1, 2, \ldots, N - 1$.

(ii) $\Rightarrow$ (i) assignment question. ■

§3 Product space

§3.1 Preparations

Given a measure space $(X, \mathcal{F}, \mu)$ with μ a σ-finite measure (i.e. there exists a sequence $\{X_n\} \subseteq \mathcal{F}$ such that $X_n \uparrow$ and $\bigcup_n X_n = X$, and $\mu(X_n) < \infty$ for each n).

↪**Definition 3.1** (Measurable): $f : X \to \overline{R}$ is $\mathcal{F}$-measurable if $\forall\, a \in \mathbb{R}, f^{-1}([-\infty, a)) \in \mathcal{F}$.

We have similar properties for f in general as in the Lebesgue setting. -For f $\mathcal{F}$-measurable, $cf, f^k, |f|, f \wedge a, f \vee b, f^+, f^-$ are all $\mathcal{F}$-measurable for $a, b, c \in \mathbb{R}$.

- For f, g $\mathcal{F}$-measurable, $f + g, f - g, f \cdot g, f \wedge g, f \vee g$ are all $\mathcal{F}$-measurable.
- If $\{f_n\}$ $\mathcal{F}$-measurable, $\sup_n f_n$, $\inf_n f_n$, $\limsup_n f_n$, $\liminf_n f_n$ are $\mathcal{F}$-measurable.

We may "dissect" functions as before. For f $\mathcal{F}$-measurable, write $f = f^+ - f^-$, and put for $n \geq 1$ and $\bullet = +, -$,

$$f_n^\bullet := \mathbb{1}_{X_n}(f^\bullet \wedge n).$$

Then, $f_n^\bullet \uparrow f^\bullet$. Put

$$\varphi_n^\bullet := \sum_{k=0}^{n2^n} \frac{k}{2^n} \mathbb{1}_{A_{n,k}^\bullet},$$

where, for $k = 0, 1, \ldots, n2^n$ for $n \geq 1$,

$$A_{n,k}^\bullet = \left\{x \in X_n : \frac{k}{2^n} \leq f_n^\bullet < \frac{k+1}{2^n}\right\} \in \mathcal{F}.$$

Then, we may define the integral of the simple function

$$\int_X \varphi_n^\bullet \, d\mu := \sum_{k=0}^{n2^n} \frac{k}{2^n} \mu(A_{n,k}^\bullet).$$

Define then

$$\int_X f^\bullet \, d\mu := \lim_n \int_X \varphi_n^\bullet \, d\mu,$$

and

$$\int_X f \, d\mu := \int_X f^+ \, d\mu - \int_X f^- \, d\mu.$$

We say, then, $f \in L^1(\mu)$ if $\int_X |f| \, d\mu < \infty$. This generalizes the notion of integration to a (slightly more) general σ-algebra.

§3.2 Product Lebesgue σ-Algebra

We will restrict our constructions to the product of 2 spaces, i.e. $\mathbb{R}^2$, but generalizes for general $\mathbb{R}^d$.

↪**Definition 3.2** (Product σ-algebra): The *product σ-algebra* of subsets of $\mathbb{R}^2$, denoted by $\mathcal{M} \otimes \mathcal{M}$ or simply $\mathcal{M}^2$, is defined as

$$\mathcal{M}^2 := \sigma(\{A \times B : A, B \in \mathcal{M}\}),$$

where

$$A \times B := \{(x, y) : x \in A, y \in B\}$$

as is standard.

Notice $\mathcal{M}^2$ contains

- rectangles $I_1 \times I_2$, I_1, I_2 intervals;
- singletons $\{(x, y)\}$;
- open sets, closed sets, and so $\mathfrak{B}(\mathbb{R}^2) := \sigma(\{\text{open sets in } \mathbb{R}^2\}) \subseteq \mathcal{M}^2$.

Given G open, then for every $x \in G$, there exists some disc centered at x contained entirely in G. Moreover, there exist $(a_1, a_2), (b_1, b_2)$ with $a_i, b_i \in \mathbb{Q}$ such that $x \in (a_1, a_2) \times (b_1, b_2) \subset G$. Then, $G = \bigcup_{x \in G} (a_1, a_2) \times (b_1, b_2)$.

↪**Definition 3.3** (Slice): Given $E \subseteq \mathbb{R}^2$, then for every $x \in R$, define

$$E_x := \{y \in \mathbb{R} : (x, y) \in E\} \subseteq \mathbb{R},$$

called the *slice* of E at x. Similarly, define for $y \in \mathbb{R}$,

$$E^y := \{x \in \mathbb{R} : (x, y) \in E\} \subseteq \mathbb{R}.$$

↪**Proposition 3.1**: If $E \in \mathcal{M}^2$, then for every $x \in \mathbb{R}$, $E_x \in \mathcal{M}$, and for every $y \in \mathbb{R}$ $E^y \in \mathcal{M}$; that is, product measurability ⇒ marginal measurability.

PROOF. Define

$$\mathcal{A} := \{E \subseteq \mathbb{R}^2 : \forall\, x \in \mathbb{R}, E_x \in \mathcal{M}\}.$$

We claim $\mathcal{A}$ a σ-algebra of subsets of $\mathbb{R}^2$.

- $\mathbb{R}^2 \in \mathcal{A}$? Yes, since for every $x \in \mathbb{R}$, $\mathbb{R}^2_x = \mathbb{R} \in \mathcal{M}$.
- Let $E \in \mathcal{A}$. Then, $E_x \in \mathcal{M}$ for every $x \in \mathbb{R}$. But we have too

$$(E^c)_x = (E_x)^c,$$

  and since $E_x \in \mathcal{M} \Rightarrow (E_x)^c \in \mathcal{M}$, it follows that $E^c \in \mathcal{A}$.
- If $\{E_n\} \subseteq \mathcal{A}$, then for every $x \in \mathbb{R}$,

$$\left(\bigcup_n E_n\right)_x = \left(\bigcup_n (E_n)_x\right) \in \mathcal{M}$$

  so $\bigcup_n E_n \in \mathcal{A}$.

Hence, $\mathcal{A}$ indeed a σ-algebra of subsets of $\mathbb{R}^2$. For every $A, B \in \mathcal{M}$, we claim $A \times B \in \mathcal{A}$. We have that for every $x \in \mathbb{R}$,

$$(A \times B)_x = \begin{cases} \varnothing \text{ if } x \notin A \\ B \text{ if } x \in A \end{cases} \in \mathcal{M},$$

hence $A \times B \in \mathcal{A}$. Thus, since such sets generate $\mathcal{M}^2$, it follows that $\mathcal{M}^2 \subseteq \mathcal{A}$, and so every set in $\mathcal{M}^2$ has the desired property.

An identical proof follows for E^y-type slices. ■

Remark 3.1: Notice we didn't prove $\mathcal{A} = \mathcal{M}^2$, indeed, because its not true.

For instance, let $E = N \times A$ with N the Vitali set and $A \in \mathcal{M}$. Then, for every $x \in A$, $E_x = \begin{cases} A \text{ if } x \in N \\ \varnothing \text{ if } x \notin N \end{cases} \in \mathcal{M}$, but $E \notin \mathcal{M}^2$, because for every $y \in \mathbb{R}$, $E^y = \begin{cases} N & \text{if } y \in A \\ \varnothing \text{ else} \end{cases}$.

In fact, there eixsts sets such that E_x and $E^y \in \mathcal{M}$ for every $x, y \in \mathbb{R}$, but $E \notin \mathcal{M}^2$ (the *Sierpinski set*).

However, if $E \subseteq \mathbb{R}^2$ a product set, i.e. $E = A \times B$ for some $A, B \subseteq \mathbb{R}$, then $A, B \in \mathcal{M} \Rightarrow E \in \mathcal{M}^2$.

↪**Definition 3.4** (Slice of sets): Let $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ a function. For every $x \in \mathbb{R}$, define

$$f_x : \mathbb{R} \to \overline{\mathbb{R}}, \qquad f_x(y) := f(x, y),$$

called the *slice* of f at x. Similarly define f^y.

⊛ **Example 3.1**: If $f = \mathbb{1}_E$ for some $E \subseteq \mathbb{R}^2$, then $f_x = \mathbb{1}_{E_x}$.

↪**Proposition 3.2**: If $f : \mathbb{R}^2 \to \overline{R}$ is $\mathcal{M}^2$-measurable, then for every $x \in \mathbb{R}$, f_x is $\mathcal{M}$-measurable, and for every $y \in \mathbb{R}$ f^y is $\mathcal{M}$-measurable.

PROOF. Observe that for every $B \subseteq \mathbb{R}$,

$$\left(f^{-1}(B)\right)_x = f_x^{-1}(B)$$

for every $x \in \mathbb{R}$, with similar for ^{y}. In particular, then, if f $\mathcal{M}^2$-measurable, then for every $a \in \mathbb{R}$, $f^{-1}([-\infty, a)) \in \mathcal{M}^2$ hence $f_x^{-1}([-\infty, a)) = (f^{-1}([-\infty, a))_x \in \mathcal{M}$, with the same idea following for ^{y}. ■

Remark 3.2:
- If $f : \mathbb{R}^2 \to R$ is continuous, then f is measurable. For every $a \in \mathbb{R}$, $f^{-1}((-\infty, a))$ open by virtue (indeed, definition) of continuity, hence in $\mathcal{M}^2$.
- If $f = \mathbb{1}_E$ for some $E \subseteq \mathbb{R}^2$, f $\mathcal{M}^2$-measurable $\Leftrightarrow$ $E \in \mathcal{M}^2$.
- In general, there exists $f : \mathbb{R}^2 \to \overline{R}$ such that f_x $\mathcal{M}$-measurable but f is not $\mathcal{M}^2$-measurable.
- If $f(x, y) = h(x)g(y)$ for some non-trivial $h, g : \mathbb{R} \to \overline{\mathbb{R}}$, then f is $\mathcal{M}^2$-measurable $\Leftrightarrow$ both h and g are $\mathcal{M}$-measurable. We show $\Leftarrow$;

$$\begin{aligned}
f^{-1}([-\infty, a)) &= \{(x, y) : h(x)g(y) < a\} \\
&= \{(x, y) : h(x) = 0, 0 < a\} \\
&\qquad \cup \left\{(x, y) : h(x) > 0, g(y) < \frac{a}{h(x)}\right\} \\
&\qquad \cup \left\{(x, y) : h(x) < 0, g(y) > \frac{a}{h(x)}\right\} \\
&= \{x : h(x) = 0\} \times \mathbb{R} \cap \{0 < a\} \qquad \in \mathcal{M}^2 \\
&\qquad \cup \left(\bigcup_{q \in \mathbb{Q}} \underbrace{\left\{x : 0 < h(x), q < \frac{a}{h(x)}\right\}}_{\in \mathcal{M}} \times \underbrace{\{y : g(y) < q\}}_{\in \mathcal{M}} \right) \\
&\qquad \cup \left(\bigcup_{q \in \mathbb{Q}} \underbrace{\left\{x : 0 > h(x), q > \frac{a}{h(x)}\right\}}_{\in \mathcal{M}} \times \underbrace{\{y : g(y) > q\}}_{\in \mathcal{M}} \right) \in \mathcal{M}^2
\end{aligned}$$

§3.3 Product Measure

↪**Definition 3.5**: Given $E \in \mathcal{M}^2$, define functions

$$I_E^{(1)} : \mathbb{R} \to \overline{\mathbb{R}}, \qquad I_E^{(1)}(x) := m(E_x)$$

and

$$I_E^{(2)} : \mathbb{R} \to \overline{\mathbb{R}}, \qquad I_E^{(2)}(y) := m(E^y).$$

↪**Theorem 3.1**: Given $E \in \mathcal{M}^2$, $I_E^{(1)}, I_E^{(2)}$ are $\mathcal{M}$-measurable functions, and in particular

$$\int_{\mathbb{R}} I_E^{(1)}(x)\,\mathrm{d}x = \int_{\mathbb{R}} I_E^{(2)}(y)\,\mathrm{d}y. \qquad \circledast$$

Proof. If indeed $I_E^{(1)}, I_E^{(2)}$ $\mathcal{M}$-measurable, then the integrals of the functions are well-defined, being non-negative functions.

Set

$$\Sigma := \left\{E \in \mathcal{M}^2 : I_{E_N}^{(1)}, I_{E_N}^{(2)} \text{ are } \mathcal{M}\text{-measurable and } \circledast \text{ holds, for } E_N := E \cap [-N,N]^2 \text{ for all } N > 0\right\}.$$

Note that for every $E \in \mathcal{M}^2$, for all $N > 0$,

$$I_{E_N}^{(1)}(x) = \begin{cases} m((E_N)_x) & \text{if } x \in [-N,N] \\ 0 & \text{o.w.} \end{cases} = \mathbb{1}_{[-N,N]}(x) I_{E_N}^{(1)}(x).$$

similarly for $I_{E_N}^{(2)}$.

Let $\mathcal{C} := \{A \times B : A, B \in \mathcal{M}\}$ (recall $\mathcal{M}^2 = \sigma(\mathcal{C})$).

- *Claim 1: $\mathcal{C} \subseteq \Sigma$*

For every $N > 0$, $E_N = (A \times B) \cap [-N,N]^2 = A_N \times B_N$ ($A_N := A \cap [-N,N]$). Then,

$$I_{E_N}^{(1)}(x) = I_{A_N \times B_N}^{(1)}(x) = \begin{cases} m(B_N) & \text{if } x \in A_N \\ 0 & \text{if } x \notin A_N \end{cases}, \qquad I_{E_N}^{(2)}(y) = I_{A_N \times B_N}^{(2)}(y) = \begin{cases} m(A_N) & \text{if } y \in B_N \\ 0 & \text{if } y \notin B_N \end{cases},$$

and so $I_{E_N}^{(1)}, I_{E_N}^{(2)}$ are measurable seeing as they are both just indicator functions of measurable sets times a constant. In particular,

$$\int_{\mathbb{R}} I_{E_N}^{(1)} = m(B_N) m(A_N) = \int_{\mathbb{R}} I_{E_N}^{(2)},$$

as required. Hence, indeed $E_N \in \Sigma$ and so $\mathcal{C} \subseteq \Sigma$.

- *Claim 2: $\mathbb{R}^2 \in \Sigma$*

For every $N > 0$,

$$I_{[-N,N]^2}^{(1)}(x) = \begin{cases} 2N & \text{if } x \in [-N,N] \\ 0 & \text{o.w.} \end{cases},$$

similar for $I_{[-N,N]^2}^{(2)}$. $I_{[-N,N]}^{2(1)}, I_{[-N,N]}^{(2)}$ are both $\mathcal{M}$-measurable, and their integrals agree, and so it follows that $R^2 \in \Sigma$.

- *Claim 3:* $E \in \Sigma \Rightarrow E^c \in \Sigma$

For each $N > 0$, denote

$$F_N := E^c \cap [-N, N]^2.$$

$I^{(1)}_{F_N} = 0$ outside of $[-N, N]$, and for $x \in [-N, N]$,

$$(F_N)_x = \left\{y : (x, y) \in E^c \cap [-N, N]^2\right\} = [-N, N] \setminus E_x = [-N, N] \setminus (E_N)_x$$

so

$$I^{(1)}_{F_N}(x) = 2N - I^{(1)}_{E_N}(x)$$

for $x \in [-N, N]$. Similarly,

$$I^{(2)}_{F_N}(y) = \begin{cases} 2N - I^{(2)}_{E_N}(y) & \text{if } y \in [-N, N] \\ 0 & \text{o.w.} \end{cases}.$$

In particular, then, $I^{(1)}_{F_N}, I^{(2)}_{F_N}$ measurable, and

$$\int_{\mathbb{R}} I^{(1)}_{F_N} = \int_{[-N,N]} 2N - I^{(1)}_{E_N} = 4N^2 - \int_{\mathbb{R}} I^{(1)}_{E_N}$$
$$\int_{\mathbb{R}} I^{(2)}_{F_N} = \int_{[-N,N]} 2N - I^{(2)}_{E_N} = 4N^2 - \int_{\mathbb{R}} I^{(2)}_{E_N},$$

but we know $\int_{\mathbb{R}} I^{(1)}_{E_N} = \int_{\mathbb{R}} I^{(2)}_{E_N}$ since $E_N \in \Sigma$, hence it follows that $\int_{\mathbb{R}} F^{(1)}_N = \int_{\mathbb{R}} F^{(2)}_N$ and so it follows that $E^c \in \Sigma$.

- *Claim 4:* $\{E_k\} \subseteq \Sigma \Rightarrow E := \bigcup_{k=1}^{\infty} E_k \in \Sigma$.

Wlog, E_n's disjoint. For $N > 0$, $E_N = \bigcup_{k=1}^{\infty} E_{k,N}$.

$$I^{(1)}_{E_N}(x) = \mathbb{1}_{[-N,N]}(x) m\left(\bigcup_{k=1}^{\infty} \left(E_{k,N}\right)_x\right) = \mathbb{1}_{[-N,N]} \sum_{k=1}^{\infty} m\left(\left(E_{k,N}\right)_x\right) = \sum_{k=1}^{\infty} I^{(1)}_{E_{k,N}}(x),$$

with similarly $I^{(2)}_{E_N}(y) = \sum_{k=1}^{\infty} I^{(2)}_{E_{k,N}}(y)$. This implies $I^{(1)}_{E_N}, I^{(2)}_{E_N}$ are $\mathcal{M}$-measurable, and in particular

$$\int_{\mathbb{R}} I^{(1)}_{E_N} = \sum_{k=1}^{\infty} \int_{\mathbb{R}} I^{(1)}_{E_{k,N}}, \qquad \int_{\mathbb{R}} I^{(2)}_{E_N} = \sum_{k=1}^{\infty} \int_{\mathbb{R}} I^{(2)}_{E_{k,N}},$$

which are equal since by assumption $E_k \in \Sigma$. Hence, $E \in \Sigma$, and thus by Claims 2-4, Σ a σ-algebra of subsets of $\mathbb{R}^2$, and thus by Claim 1 $\Sigma = \mathcal{M}^2$.

Hence, for every $E \in \mathcal{M}^2$, $E \in \Sigma$ and so all the statements hold for E_N for every $N > 0$. Then,

$$I^{(1)}_E(x) = \lim_{N\to\infty} \mathbb{1}_{[-N,N]}(x) m\left((E_N)_x\right) = \lim_{N\to\infty} m\left((E_N)_x\right) = m(E_x) = \lim_{N\to\infty} I^{(1)}_{E_N}(x),$$

and in particular $\left\{I^{(1)}_{E_N}\right\} \uparrow$, hence $I^{(1)}_E$ $\mathcal{M}$-measurable, and

$$\int_{\mathbb{R}} I^{(1)}_E = \lim_{N\to\infty} \int_{\mathbb{R}} I^{(1)}_{E_N},$$

with similar for $I_E^{(2)}$, by monotonicity. Thus, since $\int_{\mathbb{R}} I_{E_N}^{(1)} = \int_{\mathbb{R}} I_{E_N}^{(2)}$ for every N, the proof follows. ■

↪**Definition 3.6**: Define a non-negative set function on $(\mathbb{R}^2, \mathcal{M}^2)$ by

$$m(E) := \int_{\mathbb{R}} I_E^{(1)}(x)\,\mathrm{d}x = \int_{\mathbb{R}} I_E^{(2)}(x)\,\mathrm{d}x, \qquad E \in \mathcal{M}^2.$$

m is called the *Lebesgue measure on* $\mathbb{R}^2$.

↪**Proposition 3.3**: m is indeed a measure on $(\mathbb{R}^2, \mathcal{M}^2)$.

Proof.

- $m(\varnothing) = \int_{\mathbb{R}} 0 = 0$
- If $\{E_k\} \subseteq \mathcal{M}^2$ disjoint, let $E = \bigcup_{k=1}^{\infty} E_k$. Then

  $$m(E) = \sum_{k=1}^{\infty} m(E_k),$$

  since for every $x \in \mathbb{R}$, $E_x = \bigcup_{k=1}^{\infty} (E_k)_x$ disjoint, so

  $$\int_{\mathbb{R}} I_E^{(1)} = \sum_{k=1}^{\infty} \int_{\mathbb{R}} I_{E_k}^{(1)},$$

  and the proof follows.

■

Remark 3.3:
1. For any $E = I_1 \times I_2$, $m(E) = \ell(I_1) \cdot \ell(I_2)$. It follows that any singleton, and countable set, and any line on $\mathbb{R}^2$ is a null set.
2. If $A \subseteq \mathbb{R}$ is a null set in $\mathcal{M}$, then $A \times \mathbb{R}, \mathbb{R} \times A$ are null sets, in $\mathcal{M}^2$.
3. $\mathcal{M}^2$ is *not* complete under m, since for instance if $N \subset \mathbb{R}$ the Vitali set, $a \in \mathbb{R}$, then $\{a\} \times N \subseteq \{a\} \times \mathbb{R}$ is a subset of a null set, but $\{a\} \times N$ is not measurable.
4. It is possible to construct m on $\mathbb{R}^2$ through the "outer measure" approach. We take $E \subseteq \mathbb{R}^2$, and define
$$m^*(E) = \inf\left\{\sum_{n=1}^{\infty} \mathrm{Area}(R_n) : R_n\text{'s closed, finite rectangles s.t. } E \subseteq \bigcup_{n=1}^{\infty} R_n\right\}.$$
Then, m^* satisfies similar properties as the 1-dimensional analog. We then say a set E is measurable if for every $F \subseteq \mathbb{R}^2$, $m^*(F) = m^*(F \cap E) + m^*(F \cap E^c)$. Collect all such sets, $\overline{\mathcal{M}^2} := \{E \subseteq \mathbb{R}^2 : E \text{ measurable}\}$. This is a σ-algebra of subsets of $\mathbb{R}^2$, with $m := m^*|_{\overline{\mathcal{M}^2}}$ a measure when restricted to it. Indeed, m matches the Lebesgue measure defined above, and $\overline{\mathcal{M}^2}$, as suggestively notated, the completion of $\mathcal{M}^2$ under the Lebesgue measure. In addition, $\overline{\mathcal{M}^2} = \overline{\mathfrak{B}_{\mathbb{R}^2}}$.
5. The Lebesgue measure m on $\mathbb{R}^2$ is the unique measure on $\mathcal{M}^2/\mathfrak{B}_{\mathbb{R}^2}/\overline{\mathcal{M}^2}$ such that for all $I_1 \times I_2$ rectangles, $m(I_1 \times I_2) = \ell(I_1)\ell(I_2)$. This is because $\mathcal{I} := \{I_1 \times I_2 : I_1, I_2 \text{ finite intervals}\}$ is a π-system and generates $\mathfrak{B}_{\mathbb{R}^2}$.
6. The Lebesgue measure on $\mathbb{R}^2$ is translation invariant (rectangle area is invariant under translation). Namely, show that $m_Z : \mathcal{M}^2 \to [0, \infty]$, $m_z(E) := m(E + z)$ is a measure and $m_z = m$ on $\mathcal{I}$.
7. The Lebesgue measure m on $\mathbb{R}^2$ is the only measure on $\mathcal{M}^2/\mathfrak{B}_{\mathbb{R}^2}/\overline{\mathcal{M}^2}$ that is translation invariant, assigns finite values to compact sets, and assigns 1 to $[0,1] \times [0,1]$.

§3.4 Fubini's Theorem

↪**Definition 3.7**: Let $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ be $\mathcal{M}^2$-measurable and non-negative. Define the functions
$$I_f^{(1)}(x) := \int_{\mathbb{R}} f(x,y)\,\mathrm{d}y = \int_{\mathbb{R}} f_x(y)\,\mathrm{d}y, \qquad I_f^{(2)}(y) := \int_{\mathbb{R}} f(x,y)\,\mathrm{d}x = \int_{\mathbb{R}} f^y(x)\,\mathrm{d}x.$$

Remark 3.4: Given $f : \mathbb{R}^2 \to [0, \infty]$, $\mathcal{M}^2$-measurable and non-negative, the integral of f wrt the Lebesgue measure on $\mathbb{R}^2$ is denoted by $\int_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$ or $\int_{\mathbb{R}^2} f$ if there is no ambiguity.

↪**Theorem 3.2** (Tonelli's): Let $f : \mathbb{R}^2 \to [0,\infty]$ be $\mathcal{M}^2$-measurable and non-negative. Then,

$$\int_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}} I_f^{(1)}(x)\,\mathrm{d}x = \int_{\mathbb{R}} I_f^{(2)}(y)\,\mathrm{d}y,$$

or more explicitly,

$$\int_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}}\left(\int_{\mathbb{R}} f(x,y)\,\mathrm{d}y\right)\mathrm{d}x = \int_{\mathbb{R}}\left(\int_{\mathbb{R}} f(x,y)\,\mathrm{d}x\right)\mathrm{d}y.$$

Proof. Since f $\mathcal{M}^2$-measurable, non-negative, there exists $\{\varphi_n\}$-sequence of simple functions with $\varphi_n \uparrow f$, and $\int_{\mathbb{R}^2} f = \lim_n \int_{\mathbb{R}^2} \varphi_n$, where, eg

$$\varphi_n = \sum_{k=0}^{n2^n} \frac{k}{2^n}\mathbb{1}_{A_{n,k}}, \qquad A_{n,k} := \left\{(x,y) \in [-n,n]^2 : \frac{k}{2^n} \leq f(x,y) < \frac{k+1}{2^n}\right\}, k = 0,1,\dots,n2^n.$$

So,

$$\int_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \lim_n \sum_{k=0}^{n2^n} \frac{k}{2^n} m(A_{n,k}).$$

On the other hand, $\forall\, x \in \mathbb{R}$, by MON

$$\begin{aligned} I_f^{(1)}(x) = \int_{\mathbb{R}} f(x,y)\,\mathrm{d}y &= \lim_{n\to\infty} \int_{\mathbb{R}} \varphi_n(x,y)\,\mathrm{d}y \\ &= \lim_{n\to\infty} \sum_{k=0}^{n2^n} \frac{k}{2^n} I_{A_{n,k}}^{(1)}(x) \\ &= \lim_{n\to\infty} \sum_{k=0}^{n2^n} \frac{k}{2^n} m\big((A_{n,k})_x\big). \end{aligned}$$

We have then, again by MON, that

$$\int_{\mathbb{R}} I_f^{(1)}(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_{\mathbb{R}} \sum_{k=0}^{n2^n} \frac{k}{2^n} I_{A_{n,k}}^{(1)}(x)\,\mathrm{d}x = \lim_{n\to\infty} \sum_{k=0}^{n2^n} \frac{k}{2^n} \int_{\mathbb{R}} I_{A_{n,k}}^{(1)}(x)\,\mathrm{d}x.$$

Similarly, we find

$$\int_{\mathbb{R}} I_f^{(2)}(y)\,\mathrm{d}y = \lim_{n\to\infty} \sum_{k=0}^{n2^n} \frac{k}{2^n} \int_{\mathbb{R}} I_{A_{n,k}}^{(2)}(y)\,\mathrm{d}y.$$

By definition,

$$m(A_{n,k}) = \int_{\mathbb{R}} I_{A_{n,k}}^{(1)}(x)\,\mathrm{d}x = \int_{\mathbb{R}} I_{A_{n,k}}^{(2)}(y)\,\mathrm{d}y,$$

hence all of our terms actually agree, and bringing them together gives the proof. ■

↪**Definition 3.8**: Given $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ $\mathcal{M}^2$-measurable, we write $f \in L^1(\mathbb{R}^2)$ if

$$\int_{\mathbb{R}^2} |f(x,y)| \, \mathrm{d}x \, \mathrm{d}y < \infty,$$

or equivalently if

$$\int_{\mathbb{R}^2} f^+ \text{ and } \int_{\mathbb{R}^2} f^- < \infty.$$

Remark 3.5: Suppose $f \in L^1(\mathbb{R}^2)$. Then by Tonelli's,

$$\int_{\mathbb{R}^2} |f(x,y)| \, \mathrm{d}x \, \mathrm{d}y = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x,y)| \, \mathrm{d}y \right) \mathrm{d}x = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x,y)| \, \mathrm{d}x \right) \mathrm{d}y,$$

and in particular all integrals are finite; namely, $I_{|f|}^{(1)}, I_{|f|}^{(2)} \in L^1(\mathbb{R})$.

↪**Theorem 3.3** (Fubini's): If $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ is $\mathcal{M}^2$-measurable and $f \in L^1(\mathbb{R}^2)$, then
1. $I_f^{(1)}, I_f^{(2)} \in L^1(\mathbb{R})$ (product integrability $\Rightarrow$ marginal integrability)
2. $I_f^{(1)}(x)$ finite-valued for a.e. $x \in \mathbb{R} \Rightarrow f_X \in L^1(\mathbb{R})$ for a.e. $x \in \mathbb{R}$, similar for $I_f^{(2)}$, i.e. $f^y \in L^1(\mathbb{R})$ for a.e. $y \in \mathbb{R}$.
3. $\int_{\mathbb{R}^2} f(x,y) \, \mathrm{d}x \, \mathrm{d}y = \int_{\mathbb{R}} I_f^{(1)}(x) \, \mathrm{d}x = \int_{\mathbb{R}} I_f^{(2)}(y) \, \mathrm{d}y$.

Proof. Assume $f \in L^1(\mathbb{R})$. Then by Tonelli's

$$\int_{\mathbb{R}} |I_f^{(1)}| \leq \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x,y)| \, \mathrm{d}y \right) \mathrm{d}x = \int_{\mathbb{R}^2} f(x,y) \, \mathrm{d}x \, \mathrm{d}y < \infty \Rightarrow I_f^{(1)} \in L^1(\mathbb{R}).$$

We have 1. $\Rightarrow$ 2..

Now, write $f = f^+ - f^-$. $f \in L^1(\mathbb{R}^2)$ gives that $f^+, f^- \in L^1(\mathbb{R}^2)$ so f^+, f^- each finite valued a.e.. By Tonelli's, then,

$$\int_{\mathbb{R}} I_{f^+}^{(1)}(x) \, \mathrm{d}x = \int_{\mathbb{R}} I_{f^+}^{(2)}(y) \, \mathrm{d}y = \int_{\mathbb{R}^2} f^+(x,y) \, \mathrm{d}x \, \mathrm{d}y < \infty,$$

same with f^-. Then, $I_{f^+}^{(1)}, I_{f^+}^{(2)}, I_{f^-}^{(1)}, I_{f^-}^{(2)} \in L^1(\mathbb{R})$, hence are finite-valued a.e.. By linearity on L^1 functions, then

$$\int_{\mathbb{R}} I_{f^+}^{(1)}(x) \, \mathrm{d}x - \int_{\mathbb{R}} I_{f^-}^{(1)}(x) \, \mathrm{d}x = \int_{\mathbb{R}} I_{f^+}^{(1)} - I_{f^-}^{(1)}.$$

For a.e. $x \in \mathbb{R}$, $f_x^+, f_x^- \in L^1(\mathbb{R})$, so by linearity

$$I_{f^+}^{(1)}(x) - I_{f^-}^{(1)}(x) = \int_{\mathbb{R}} f_x^+(y) \, \mathrm{d}y - \int_{\mathbb{R}} f_x^-(y) \, \mathrm{d}y = \int_{\mathbb{R}} (f_x^+ - f_x^-) = \int_{\mathbb{R}} f_x,$$

so

$$\int_{\mathbb{R}} I_{f^+}^{(1)} - \int_{\mathbb{R}} I_{f^-}^{(1)} = \int_{\mathbb{R}} I_f^{(1)},$$

with similarly for

$$\int_{\mathbb{R}} I_{f^+}^{(2)} - \int_{\mathbb{R}} I_{f^-}^{(2)} = \int_{\mathbb{R}} I_f^{(2)}.$$

All together, then,

$$\int_{\mathbb{R}} I_f^{(1)} = \int_{\mathbb{R}} I_f^{(2)} = \int_{\mathbb{R}^2} (f^+ - f^-) = \int_{\mathbb{R}^2} f.$$

■

Remark 3.6: In general, $I_f^{(1)}, I_f^{(2)} \in L^1(\mathbb{R}) \nRightarrow f \in L^1(\mathbb{R}^2)$. For instance, let

$$f(x,y) = \begin{cases} 1 & \text{if } x < y < x+1 \\ -1 & \text{if } x - 1 < y < x. \\ 0 & \text{else} \end{cases}$$

Then, for all $x \in \mathbb{R}$,

$$I_f^{(1)}(x) = \int_{\mathbb{R}} f(x,y) \, \mathrm{d}y = 0,$$

and similarly

$$I_f^{(2)}(y) = \int_{\mathbb{R}} f(x,y) \, \mathrm{d}x = 0,$$

so $I_f^{(1)}, I_f^{(2)} \in L^1(\mathbb{R})$, but $f \notin L^1(\mathbb{R}^2)$.

Remark 3.7: If $f : \mathbb{R}^2 \to \overline{R}$ is $\overline{\mathcal{M}^2}$-measurable, Tonelli's, Fubini's still hold. We'll use the same notations in this case.

In fact, there exists a $\tilde{f} : \mathbb{R}^2 \to \overline{R}$ that is $\mathcal{M}^2$-measurable such that $\tilde{f} = f$ a.e. (exercise).

Remark 3.8: The constructions above extend to $\mathbb{R}^d$, $d \geq 3$. In particular, we have

- $\mathcal{M}^d := \sigma(\{A_1 \times \cdots \times A_d : A_i \in \mathcal{M}\})$.
- The product measure m is the Lebesgue measure on $(\mathbb{R}^d, \mathcal{M}^d)$.
- $\overline{\mathcal{M}^d}$ is the completion of $\mathcal{M}^d$ under m.
- Tonelli's, Fubini's hold, with "d-embedded" integrals.
- m shares similar properties on $\mathbb{R}^d$ as on $\mathbb{R}$;
  - translation invariance,
  - scaling property,
  - regularity, (outer: for every $E \in \mathcal{M}^d$, $m(E) = \inf\{m(G) : G \text{ open s.t. } E \subseteq G\}$, inner: for every $E \in \mathcal{M}^d$, $m(E) = \sup\{m(K) : K \text{ compact s.t. } E \supseteq K\}$).

§4 Differentiation

In the Riemann setting, differentiation and integration are closely related. For instance, if $F(x) := \int_a^x f(t) \, \mathrm{d}t$ for some Riemann-integrable f on $[a,b]$ and $x \in [a,b]$, then F is differentiable and $F' = f$ on (a,b). Or, if F differentiable, and F' is Riemann integrable on some $[a,b]$, then $F(b) - F(a) = \int_a^b F'(t) \, \mathrm{d}t$. How much does this extend to the Lebesgue setting?

§4.1 Hardy-Littlewood Maximal Function

↪**Definition 4.1** (Hardy-Littlewood Maximal Function)**:** Suppose $f \in L^1(\mathbb{R})$. The *Hardy-Littlewood Maximal Function* (H-L max.), denoted f^*, is defined as

$$f^*(x) := \sup_{I \in \mathcal{J}(x)} \frac{1}{m(I)} \int_I |f|,$$

where $\mathcal{J}(x) := \{I : I \text{ an open interval containing } x\}$.

↪**Proposition 4.1:** Given $f \in L^1(\mathbb{R})$, f^* is measurable.

Proof. $f^* \geq 0$, so it suffices to show that for every $a \geq 0$, $\{f^* > a\}$ is measurable. Let $x \in \{f^* > a\}$. Then, $a < f^*(x)$, hence there must exist some $I \in \mathcal{J}(x)$ such that $\frac{1}{m(I)} \int_I |f| > a$. I is open, and $x \in I$, so there exists some $\delta > 0$ such that $(x - \delta, x + \delta) \subseteq I$. For every $y \in (x - \delta, x + \delta)$, $y \in I$, hence $I \in \mathcal{J}(y)$. So, $f^*(y) \geq \frac{1}{m(I)} \int_I |f| > a$. Thus, $y \in \{f^* > a\}$ as well. It follows, then, that $(x - \delta, x + \delta) \subseteq \{f^* > a\}$, hence $\{f^* > a\}$ is open, and so in particular is measurable. ■

↪**Lemma 4.1** (Vitali's Covering Lemma)**:** Assume that $\mathcal{J} := \{I_1, ..., I_N\}$ a finite collection of open intervals. Then, there exists a sub-collection $\{I_{k_1}, ..., I_{k_M}\} \subset \mathcal{J}$ such that $I_{k_i} \cap I_{k_j} = \varnothing$ for all $i \neq j$ and

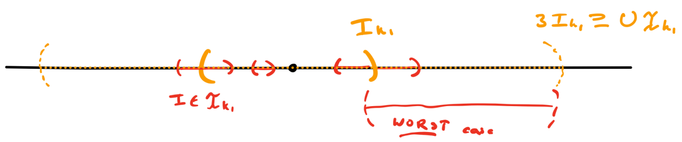
$$m\left(\bigcup_{i=1}^N I_i\right) \leq 3 \sum_{j=1}^M m\left(I_{k_j}\right).$$

Proof. Assume wlog that $m(I_i) < \infty$ for all $1 \leq i \leq N$; if otherwise exists an i such that $m(I_i) = \infty$, then simply take your subcollection $I_{k_1} := I_i$, and the claim holds trivially.

Begin with the largest interval in $\mathcal{J}$, call it I_{k_1}. Let

$$\mathcal{J}_{k_1} := \left\{I \in \mathcal{J} : I \cap I_{k_1} \neq \varnothing\right\}.$$

For any $I \in \mathcal{J}_{k_1}$, $I \cap I_{k_1} \neq \varnothing$ and $m(I) \leq m\left(I_{k_1}\right)$, so in particular $I \subseteq 3I_{k_1}$ (if $I_{k_1} = (a, b)$, $3I_{k_1} := (a - 3(b - a), a + 3(b - a))$).



Then, in particular

$$\bigcup_{I \in \mathcal{J}_{k_1}} I \subseteq 3I_{k_1}.$$

Consider now $\mathcal{I} \setminus \mathcal{I}_{k_1}$, and choose the largest interval in the remaining part of the collection. Call it I_{k_2}. Set

$$\mathcal{I}_{k_2} := \left\{I \in \mathcal{I} \setminus \mathcal{I}_{k_1} : I \cap I_{k_2} \neq \varnothing\right\}.$$

Similarly to before, $\bigcup_{I \in \mathcal{I}_{k_2}} I \subseteq 3I_{k_2}$. By choice, too, $I_{k_1} \cap I_{k_2} = \varnothing$.

Repeat this process, until $\mathcal{I} \setminus \left(\mathcal{I}_{k_1} \cup \cdots \cup \mathcal{I}_{k_M}\right) = \varnothing$, i.e. we have no intervals left in the original collection. Then, we obtain $I_{k_1}, \ldots, I_{k_M}$ disjoint intervals, with corresponding subcollections $\mathcal{I}_{k_1}, \ldots, \mathcal{I}_{k_M}$ forming a partition of $\mathcal{I}$. Then,

$$m\left(\bigcup_{n=1}^{N} I_n\right) = \sum_{j=1}^{M} m\left(\bigcup_{I \in \mathcal{I}_{k_j}} I\right) \leq 3\sum_{j=1}^{M} m\left(I_{k_j}\right).$$

■

↪**Proposition 4.2**: Suppose $f \in L^1(R)$ and let f^* be the H-L max function of f. Then, for every $\varepsilon > 0$,

$$m(\{x \in \mathbb{R} : f^*(x) > \varepsilon\}) \leq \frac{3}{\varepsilon}\|f\|_1 = \frac{3}{\varepsilon}\int_{\mathbb{R}} |f|.$$

Proof. Fix $\varepsilon > 0$ and put $B := \{x \in \mathbb{R} : f^*(x) > \varepsilon\}$. By inner regularity,

$$m(B) = \sup\{m(K) : B \supseteq K \text{ compact}\}.$$

It suffices to show then that $m(K) \leq \frac{3}{\varepsilon}\|f\|_1$ for every compact $K \subseteq B$. For every $x \in K$, $f^*(x) > \varepsilon$ so there exists some open interval I_x such that $x \in I_x$ and $\frac{1}{m(I_x)}\int_{I_x} |f| > \varepsilon$. Hence, we may cover $K \subseteq \bigcup_{x \in K} I_x$. Since K compact it admits a finite subinterval, call it $\mathcal{I} = \{I_1, \ldots, I_N\}$, such that $K \subseteq \bigcup_{n=1}^{N} I_n$. By the Covering Lemma,

$$m(K) \leq m\left(\bigcup_{n=1}^{N} I_n\right) \leq 3\sum_{j=1}^{M} m\left(I_{k_j}\right),$$

for some disjoint subcollection $I_{k_1}, \ldots, I_{k_M}$. Meanwhile, for every $1 \leq j \leq M$,

$$m\left(I_{k_j}\right) < \frac{1}{\varepsilon}\int_{I_{k_j}} |f|,$$

hence, we find

$$m(K) \leq 3\sum_{j=1}^{M} \frac{1}{\varepsilon}\int_{I_{k_j}} |f| = \frac{3}{\varepsilon}\int_{\bigcup_{j=1}^{M} I_{k_j}} |f| \leq \frac{3}{\varepsilon}\int_{\mathbb{R}} |f| = \frac{3}{\varepsilon}\|f\|_1.$$

■

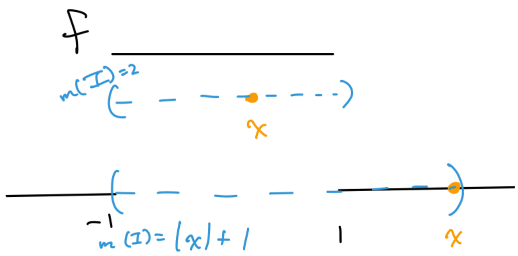
↪**Corollary 4.1**: Given $f \in L^1(\mathbb{R})$, f^* is finite-valued a.e..

Proof. For every $N > 0$, $m(\{f^* > N\}) \leq \frac{3}{N}\|f\|_1$. Taking $N \to \infty$, we find then $m(\{f^* > N\}) \to 0$, and since $m(\{f^* = \infty\}) \leq m(\{f^* > N\})\ \forall\, N > 0$ it follows that $m(\{f^* = \infty\}) = 0$. ■

Remark 4.1: While a Markov-like inequality, f^* need not be integrable in general. For instance, let $f = \mathbb{1}_{[-1,1]} \in L^1(\mathbb{R})$. Then, consider f^*, and in particular consider the average of f over intervals $I = (a, b)$;

$$\frac{1}{b-a}\int_a^b f(t)\,\mathrm{d}t = \begin{cases} 0 & \text{if } (a,b) \cap [-1,1] = \varnothing \\ \frac{\min\{b,1\} - \max\{a,-1\}}{b-a} & \text{if } (a,b) \cap [-1,1] \neq \varnothing. \end{cases}.$$

So, we find that $f^*(x) = 1$ if $x \in (-1,1)$ (take your $I = [-1,1]$, this achieves max), and $f^*(x) = \frac{2}{|x|+1}$ if $x \notin (-1,1)$ (you want as much of the $[-1,1]$ support as possible, and with your other endpoint as close to x as possible). f^* not integrable.



§4.2 Lebesgue Differentiation Theorem

↪**Theorem 4.1** (Lebesgue Differentiation Theorem): Given $f \in L^1(\mathbb{R})$, for a.e. $x \in \mathbb{R}$, if $\{I_n\}$ a sequence of open intervals such that $x \in I_n\ \forall\, n \geq 1$ and $\lim_{n\to\infty} m(I_n) = 0$ (we say $\{I_n\}$ a sequence of intervals *shrinking to* x), then

$$\lim_{n\to\infty} \frac{1}{m(I_n)} \int_{I_n} |f(t) - f(x)|\,\mathrm{d}t = 0.$$

In particular,

$$\lim_{n\to\infty} \frac{1}{m(I_n)} \int_{I_n} f(t)\,\mathrm{d}t = f(x).$$

Proof. The "In particular" comes from the fact that, for x such that $f(x) < \infty$,

$$\left|\frac{1}{m(I_n)}\int_{I_n} f(t)\,\mathrm{d}t - f(x)\right| = \left|\frac{1}{m(I_n)}\int_{I_n} f(t) - f(x)\,\mathrm{d}t\right| \leq \frac{1}{m(I_n)}\int_{I_n} |f(t) - f(x)|\,\mathrm{d}t,$$

so if the RHS $\to 0$, so does the left.

Without loss of generality, assume f finite valued everywhere, and only use finite-valued intervals I_n. For every $k \geq 1$, define

$$B_k := \left\{ x \in \mathbb{R} : \exists \{I_n\} \subseteq \mathcal{I}(x) \text{ with } \lim_{n\to\infty} m(I_n) = 0 \text{ s.t.} \limsup_{n\to\infty} \frac{1}{m(I_n)} \int_{I_n} |f(t) - f(x)| \,\mathrm{d}t \geq \frac{1}{k} \right\}.$$

Notice $B_k \uparrow$ and $\bigcup_{k=1}^{\infty} B_k = \{x \in \mathbb{R} : \text{theorem fails}\}$. So, it suffices to show that $m(B_k) = 0$ for every $k \geq 1$.

Fix an arbitrary $\varepsilon > 0$. Continuous, compactly supported functions are dense in $L^1(\mathbb{R})$ so we may find $g \in C_c(\mathbb{R})$ such that $\|f - g\|_1 \leq \varepsilon$. Since g continuous and compactly supported, for every $x \in \mathbb{R}$ and $k \geq 1$, there exists some $\alpha > 0$ such that if $|t - x| \leq \alpha$, $|g(t) - g(x)| \leq \frac{1}{3k}$.

Given any $x \in \mathbb{R}$ and any sequence $\{I_n\} \subseteq \mathcal{I}(x)$ with $\lim_n m(I_n) = 0$, we have

$$\begin{aligned} \frac{1}{m(I_n)} \int_{I_n} |f(t) - f(x)| \,\mathrm{d}t \leq \frac{1}{m(I_n)} \int_{I_n} |f(t) - g(t)| \,\mathrm{d}t \quad &(1) \\ + \frac{1}{m(I_n)} \int_{I_n} |g(t) - g(x)| \,\mathrm{d}t \quad &(2) \\ + |g(x) - f(x)| \quad &(3) \end{aligned}$$

by triangle inequality, adding/subtracting $g(t), g(x)$. We know that when n sufficiently large such that $m(I_n) < \alpha$, $|g(t) - g(x)| \leq \frac{1}{3k}\ \forall\, t \in I_n$, hence (2) $\leq \frac{1}{3k}$ for sufficiently large n. For x to be in B_k, we need too that $\limsup_n((1) + (2) + (3)) > \frac{1}{k}$. But we know that (2) $\leq \frac{1}{3k}$ for all sufficiently large n, we must have that $\limsup_n((1) + (3)) > \frac{2}{3k}$. Let

$$C_k := \left\{ x \in \mathbb{R} : \limsup_n (1) > \frac{1}{3k} \right\}, \qquad D_k := \left\{ x \in \mathbb{R} : \limsup_n (3) > \frac{1}{3k} \right\},$$

then remark $m(B_k) \leq m(C_k) + m(D_k)$ since $B_k \subseteq C_k \cup D_k$. Then,

$$m(D_k) = m\left(\left\{ |f - g| > \frac{1}{3k} \right\} \right) \leq 3k \, \|f - g\|_1 \leq 3k\varepsilon,$$

by Markov's, and

$$\begin{aligned} m(C_k) &= m\left(\left\{ \limsup_n \frac{1}{m(I_n)} \int (I_n) \, |f - g| > \frac{1}{3k} \right\} \right) \\ &\leq m\left(\left\{ (f - g)^* > \frac{1}{3k} \right\} \right) \leq 3 \cdot 3k \, \|f - g\|_1 = 9k\varepsilon, \end{aligned}$$

by using the previous H-L inequality. Hence, we find

$$m(B_k) \leq 12k\varepsilon,$$

and, sending $\varepsilon \to 0$ we find $m(B_k) = 0$, completing the proof. ■

↪**Definition 4.2** (Lebesgue Point): We call x a *Lebesgue point* of f if the Lebesgue Differentiation Theorem holds for f at x.

Remark 4.2: In the statement, the "a.e. $x \in \mathbb{R}$" cannot be replaced with "pointwise". For example, consider $f = \mathbb{1}_{[0,1]}$ and $I_n = \left(-\frac{1}{n}, \frac{1}{n}\right) \in \mathcal{I}(0)$. Then

$$\frac{1}{m(I_n)} \int_{I_n} f = \frac{2}{n} \int_{-\frac{1}{n}}^{\frac{1}{n}} \mathbb{1}_{[0,1]} = \frac{1}{2} \neq f(0),$$

hence 0 not a Lebesgue point of f.

↪**Corollary 4.2**: If $f \in L^1(\mathbb{R})$, then for a.e. $x \in \mathbb{R}$,

$$\begin{aligned} 0 &= \lim_{h \to 0} \frac{1}{2h} \int_{x-h}^{x+h} |f(t) - f(x)| \, \mathrm{d}t \qquad \text{(i)} \\ &= \lim_{h \to 0} \frac{1}{2h} \int_{-h}^{h} |f(x+y) - f(x)| \, \mathrm{d}y \qquad \text{(ii)}. \end{aligned}$$

PROOF. (i) $\Rightarrow$ (ii) by translation of the integral.

If $x \in \mathbb{R}$ is such that (i) fails, then there is a sequence $\{h_n\} \in \mathbb{R}^+$ such that $\lim h_n = 0$ and $\lim_n \frac{1}{2h_n} \int_{x-h_n}^{x+h_n} |f(t) - f(x)| \, \mathrm{d}t \neq 0$. Then, the Lebesgue Diff Thm fails at x for $I_n = (x - h_n, x + h_n)$ so x not a Lebesgue point. Hence, $\{x : \text{(i) fails}\}$ a null set.

In particular, this implies that $\lim_{h \to 0} \frac{1}{2h} \int_{x-h}^{x+h} f(t) \, \mathrm{d}t = f(x)$ for a.e. $x \in \mathbb{R}$, so as a function of x, $\lim_{h \to 0} \frac{1}{2h} \int_{x-h}^{x+h} f(t) \, \mathrm{d}t$ measurable. ■

↪**Corollary 4.3**: Given $f \in L^1(\mathbb{R})$, then $|f| \leq f^*$ a.e..

PROOF. Apply LDT to $|f|$. This implies that for a.e. $x \in \mathbb{R}$, $|f(x)| = \lim_{n \to \infty} \frac{1}{m(I_n)} \int_{I_n} |f(t)| \, \mathrm{d}t$, $I_n := \left(x - \frac{1}{n}, x + \frac{1}{n}\right)$. By definition, $\lim_{n \to \infty} \frac{1}{m(I_n)} \int_{I_n} |f(t)| \, \mathrm{d}t \leq f^*(x)$, being the supremum of such quantities, and the proof follows. ■

↪**Theorem 4.2**: Given $f \in L^1(\mathbb{R})$ and $a \in \mathbb{R}$, define $F(x) := \int_a^x f(t) \, \mathrm{d}t$ for every $x \geq a$. Then, F is uniformly continuous, $F'(x)$ exists and is equal to $f(x)$ for a.e. $x \in \mathbb{R}$.

↪**Lemma 4.2**: If $f \in L^1(\mathbb{R})$, then for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\forall$ I-interval with $m(I) \leq \delta$, $\int_I |f| \leq \varepsilon$.

PROOF. Recall $f^* < \infty$ a.e.. Let $A_N := \{f^* > N\}$, then $\mathbb{1}_{A_N} \to 0$ almost everywhere as $N \to \infty$. Hence,

$$\lim_{N \to \infty} \int_{\mathbb{R}} \mathbb{1}_{A_N} |f| = 0$$

by DOM, and so

$$\int_{\{f^*>N\}} |f| \to 0$$

as $N \to \infty$. Given ε, then, $\exists N$ such that $\int_{\{f^*>N\}} |f| \leq \frac{\varepsilon}{2}$. Let $\delta = \frac{\varepsilon}{2N}$. Then, for every I (wlog open) with $m(I) < \delta$,

$$\begin{aligned}
\int_I |f| &= \int_{I\cap\{f^*>N\}} |f| + \int_{I\cap\{f^*\leq N\}} |f| \\
&\leq \int_{\{f^*>N\}} |f| + \int_{I\cap\{f^*\leq N\}} |f| \\
&\leq \frac{\varepsilon}{2} + N \cdot m(I) \\
&\leq \frac{\varepsilon}{2} + N\frac{\varepsilon}{2N} = \varepsilon.
\end{aligned}$$

■

Proof. (of ↪Theorem 4.2) For every $\varepsilon > 0$, let $\delta > 0$ as in the lemma. Then for every $x > y \geq a$ such that $|x - y| \leq \delta$,

$$|F(x) - F(y)| = |\int_y^x f(t)\, \mathrm{d}t| \leq \int_{(y,x)} |f| \leq \varepsilon,$$

so F uniformly continuous.

Let $x \in \mathbb{R}$ be a Lebesgue point of f and such that $f(x)$ is finite valued. Then,

$$\begin{aligned}
\left|\frac{1}{h}(F(x+h) - F(x)) - f(x)\right| &= \left|\frac{1}{h}\int_x^{x+h} f(t)\, \mathrm{d}t - f(x)\right| \\
&= \left|\frac{1}{h}\int_x^{x+h} f(t) - f(x)\, \mathrm{d}t\right| \\
&\leq \frac{1}{h}\int_x^{x+h} |f(t) - f(x)|\, \mathrm{d}t \\
&\leq 2\frac{1}{2h}\int_{x-h}^{x+h} |f(t) - f(x)|\, \mathrm{d}x \to 0 \text{ as } h \to 0^+.
\end{aligned}$$

We can similarly show that $\lim_{h\to 0}\left|\frac{1}{h}(F(x) - F(x-h)) - f(x)\right| = 0$. ■

Remark 4.3: In general, the a.e. statement cannot be dropped. For instance, if $f = \mathbb{1}_{\{0\}}$, $F \equiv 0$ and $F' \equiv 0$ but $F'(0) \neq f(0)$.

§4.3 Monotonic (Increasing) Functions

Let F an increasing function on $[a, b]$ (we restrict fo finite-valued functions). If needed, we can extend F to beyond this interval by setting $F(x) = F(a)$ everywhere to the left of a, and $F(x) = F(b)$ everywhere to the right of b; then F still increasing on $\mathbb{R}$.

↪**Proposition 4.3**: If F increasing on $[a, b]$, then F is continuous except at most countably many points in $[a, b]$. In particular, F is measurable.

Remark 4.4: For general functions and $x \in \mathbb{R}$, we define

$$\overline{D_r}F(x) := \limsup_{h \to 0} \frac{F(x+h) - F(x)}{h}, \qquad \underline{D_r}F(x) := \liminf_{h \to 0} \frac{F(x+h) - F(x)}{h}$$

$$\overline{D_\ell}F(x) := \limsup_{h \to 0} \frac{F(x) - F(x-h)}{h}, \qquad \underline{D_\ell}F(x) := \liminf_{h \to 0} \frac{F(x) - F(x-h)}{h}.$$

If $\overline{D_r}F(x) = \underline{D_r}F(x) = \overline{D_\ell}F(x) = \underline{D_\ell}F(x)$, $F'(x)$ exists and equals this equal limit.

↪**Proposition 4.4**: Assume F increasing on $[a,b]$. Then, F' exists a.e. on $[a,b]$.

Proof. For every $x \in [a,b]$, $\underline{D_r}F(x) \leq \overline{D_r}F(x), \underline{D_\ell}F(x) \leq \overline{D_\ell}F(x)$. If we can show $\overline{D_r}F(x) \leq \underline{D_\ell}F(x)$ and $\overline{D_\ell}F(x) \leq \underline{D_r}F(x)$ a.e., we'd be done.

Set $E := \left\{x \in [a,b] : \overline{D_\ell}F(x) > \underline{D_r}F(x)\right\}$. We want to show $m(E) = 0$. For $p < q \in \mathbb{Q}$, set

$$E_{pq} := \left\{x \in [a,b] : \overline{D_\ell}F(x) > q \text{ and } \underline{D_r}F(x) < p\right\}.$$

Notice that $E = \bigcup_{p<q\in\mathbb{Q}} E_{p,q}$, which is a countable union, so it suffices to show that $m\left(E_{p,q}\right) = 0$.

Suppose otherwise, that there is some $\delta > 0$ such that $m\left(E_{p,q}\right) = \delta$. fix any $0 < \varepsilon < \frac{\delta}{2}$. Choose an open $G \supseteq E_{p,q}$ with $m(G) \leq m\left(E_{p,q}\right) + \varepsilon = \delta + \varepsilon$. Consider

$$\mathcal{J} := \left\{I = [x, x+h] \subseteq G : \frac{F(x+h) - F(x)}{h} < p\right\}.$$

For every $x \in E_{p,q}$, $x \in G$-open, and $\underline{D_r}F(x) < p$. So, for every $x \in E_{p,q}$, there exists $I = [x, x+h] \in \mathcal{J}$ for arbitrarly small h. In particular, $\mathcal{J}$ a *Vitali covering* of $E_{p,q}$ (see following lemma). Hence, there exists a disjoint subcollection $I_1, ..., I_N \in \mathcal{J}$ such that $m\left(E_{p,q} \setminus \bigcup_{i=1}^N I_i\right) \leq \varepsilon$. Write $I_i = [x_i, x_i + h_i], i = 1, ..., N$. Define

$$\tilde{G} := \bigcup_{i=1}^N (x_i, x_i + h_i), \qquad \widetilde{E_{p,q}} = E_{p,q} \cap \tilde{G}.$$

Then,

$$m\left(\widetilde{E_{p,q}}\right) = m\left(E_{p,q} \cap \tilde{G}\right) = m\left(E_{p,q} \cap \bigcup_{i=1}^N I_n\right) = m\left(E_{p,q}\right) - m\left(E_{p,q} \setminus \bigcup_{i=1}^N I_i\right) \geq \delta - \varepsilon.$$

Since $\widetilde{E_{p,q}} \subseteq \tilde{G}$ and $\tilde{G}$ is open, define $\tilde{\mathcal{J}} := \left\{\tilde{I} = [y-r, y] \subseteq \tilde{G} : \frac{F(y) - F(y-r)}{r} > q\right\}$. Then, $\tilde{\mathcal{J}}$ is a Vitali covering of $\widetilde{E_{p,q}}$, and we can extract disjoint $\tilde{I}_1, ..., \tilde{I}_M \in \tilde{\mathcal{J}}$ such that $m\left(\widetilde{E_{p,q}} \setminus \bigcup_{j=1}^M \tilde{I}_j\right) \leq \varepsilon$. Hence,

$$\begin{aligned} m\left(\bigcup_{j=1}^M \tilde{I}_j\right) &= m\left(\widetilde{E_{p,q}}\right) - m\left(\widetilde{E_{p,q}} \setminus \bigcup_{j=1}^M \tilde{I}_j\right) \\ &\geq \delta - \varepsilon - \varepsilon = \delta - 2\varepsilon. \end{aligned}$$

Write $\tilde{I}_j := \left[y_j - r_j, y_j\right], j = 1, ..., M$. Then, $\tilde{I}_1, ..., \widetilde{I_M}$ form a disjoint subintervals of $I_1, ..., I_N$. Since F is increasing,

$$\sum_{j=1}^{M}\left(F\left(y_j\right) - F\left(y_j - r_j\right)\right) \leq \sum_{i=1}^{N}(F(x_i + h) - F(x_i)).$$

We have that

$$\text{RHS} \leq p\sum_{i=1}^{N} h_i = pm\left(\bigcup_{i=1}^{N} I_i\right) \leq pm(G) \leq p(\delta + \varepsilon),$$

and similarly,

$$\text{LHS} \geq q\sum_{j=1}^{M} r_j = qm\left(\bigcup_{j=1}^{M} \tilde{I}_j\right) \geq q(\delta - 2\varepsilon),$$

hence we find

$$q(\delta - 2\varepsilon) \leq p(\delta + \varepsilon),$$

but ε arbitrary, so we find $q\delta \leq p\delta$, contradicting $p < q$, hence $\delta = 0$. ■

Remark 4.5: We tacitly used that $\overline{D_r}F$, etc, are measurable to say that E measurable. This needs to be proven.

↪**Lemma 4.3** (Vitali's Covering Theorem): Given a set $E \subseteq \mathbb{R}$, a collection $\mathcal{J}$ of intervals is called a *Vitali covering* of E if for every $x \in E$ and $\varepsilon > 0$, there is an $I \in \mathcal{J}$ such that $x \in I$ and $m(I) < \varepsilon$.

If $E \in \mathcal{M}$ with $m(E) < \infty$ and $\mathcal{J}$ a Vitali covering of E, then for every $\varepsilon > 0$, there is a finite subcollection $I_1, I_2, ..., I_N \in \mathcal{J}$ such that $I_i \cap I_j = \varnothing$ for $i \neq j$, and

$$m\left(E \setminus \bigcup_{i=1}^{N} I_i\right) \leq \varepsilon.$$

Proof. Assume, wlog, $m\left(\bigcup_{I\in\mathcal{J}} I\right) < \infty$. Else, let G open such that $G \supseteq E$ and $m(G) < \infty$, and redefine $\mathcal{J}' := \left\{I \in \mathcal{J} : I \subseteq G\right\}$, then, if $\mathcal{J}$ a Vitali covering of E, so is $\mathcal{J}'$. Defining then

$$\alpha_1 := \sup\left\{m(I) : I \in \mathcal{J}\right\},$$

we know $\alpha_1 < \infty$. Then, $\exists\, I_1 \in \mathcal{J}$ such that $m(I_1) > \frac{\alpha_1}{2}$. Then, consider

$$\mathcal{J}_1 := \left\{I \in \mathcal{J} : I \cap I_1 = \varnothing\right\}.$$

Define

$$\alpha_2 := \sup\left\{m(I) : I \in \mathcal{J}_1\right\} < \infty,$$

so there exists $I_2 \in \mathcal{J}_1$ such that $m(I_2) > \frac{\alpha_2}{2}$, and put $\mathcal{J}_2 := \{I \in \mathcal{J} : I \cap I_1 = \varnothing \text{ and } I \cap I_2 = \varnothing\}$. Repeat this procedure; this will generate a sequence $\{\alpha_k\}$ and $\{I_k\}$ such that the I_k's are disjoint, and $\alpha_k = \sup\{m(I) : I \in \mathcal{J}, I \cap I_j = \varnothing \; \forall j = 1, \dots, k-1\}$. Since $\bigcup_{k=1}^{\infty} I_k \subseteq \bigcup_{I \in \mathcal{J}} I$ and disjointness,

$$\sum_{k=1}^{\infty} m(I_k) = m\left(\bigcup_{k=1}^{\infty} I_k\right) < \infty.$$

In addition, $m(I_k) > \frac{\alpha_k}{2}$ hence $\sum_{k=1}^{\infty} m(I_k) \geq \sum_{k=1}^{\infty} \frac{\alpha_k}{2}$, noticing that $\alpha_k \to 0$ as $k \to \infty$ and in particular this means $\sum_{k=1}^{\infty} \frac{\alpha_k}{2}$ a converging series. Fix $\varepsilon > 0$, then there exists some N sufficiently large such that

$$\sum_{k=N+1}^{\infty} m(I_k) < \frac{\varepsilon}{5}.$$

We claim that $\{I_i\}_{i=1}^{N}$ satisfies our desired properties, namely that $m\left(E \setminus \bigcup_{i=1}^{N} I_i\right) < \varepsilon$. It suffices to show $m\left(E \setminus \bigcup_{i=1}^{N} \overline{I_i}\right) < \varepsilon$, since this adds/removes only points so doesn't change the measure. Since $\bigcup_{i=1}^{N} \overline{I_i}$ closed, then for every $x \in E \setminus \bigcup_{i=1}^{N} \overline{I_i}$, $\operatorname{dist}\left(x, \bigcup_{i=1}^{N} \overline{I_i}\right) = \lambda > 0$. Since $\mathcal{J}$ a Vitali covering, there is some $I^* \in \mathcal{J}$ such that $x \in I^*$ and $m(I^*) < \lambda$. Hence, it must be that $I^* \cap I_i = \varnothing$ for every $i = 1, \dots, N$. Hence, $m(I^*) \leq \alpha_{N+1}$. Let $N^* \geq N+1$ be such that $\alpha_{N^*+1} < m(I^*) \leq \alpha_{N^*}$. So, there is a $j = N+1, \dots, N^*$ such that $I^* \cap I_j \neq \varnothing$ (we "start seeing" I^* at the jth step). Now,

$$m(I^*) \leq \alpha_{N^*} \leq \alpha_j \leq 2m\left(I_j\right).$$

In particular, $I^* \subseteq$ "$5I_j$", where $5I_j$ is I_j "expanded" 5 times. So,

$$E \setminus \bigcup_{i=1}^{N} \overline{I_i} \subseteq \bigcup_{k=N+1}^{\infty} \text{"}5I_k\text{"},$$

so

$$m\left(E \setminus \bigcup_{i=1}^{N} \overline{I_i}\right) \leq m\left(\bigcup_{k=N+1}^{\infty} \text{"}5I_k\text{"}\right) \leq 5 \sum_{k=N+1}^{\infty} m(I_k) = 5 \cdot \frac{\varepsilon}{5} = \varepsilon,$$

as we aimed to show. ■

↪**Proposition 4.5**: Assume $F : [a, b] \to \mathbb{R}$ is increasing. Then, $F' \in L^1([a, b])$, $F' \geq 0$ a.e. on $[a, b]$, and $\int_a^b F'(t)\, dt \leq F(b) - F(a)$.

PROOF. $F' \geq 0$ clear. For a.e. $x \in [a, b]$, $F'(x) = \lim_{n \to \infty} G_n(x)$ where $G_n(x) := \frac{F\left(x + \frac{1}{n}\right) - F(x)}{\frac{1}{n}}$, expanding F to be constant its last value outside of $[a, b]$ if necessary. Then, by Fatou,

$$
\begin{aligned}
\int_a^b F'(x)\,\mathrm{d}x &\leq \liminf_n \int_a^b G_n(x)\,\mathrm{d}x \\
&= \liminf_n \left(n \left(\int_a^b F\left(x + \frac{1}{n} \right) \mathrm{d}x - \int_a^b F(x)\,\mathrm{d}x \right) \right) \\
&= \liminf_n \left(n \left(\int_{a+\frac{1}{n}}^{b+\frac{1}{n}} F(t)\,\mathrm{d}t - \int_a^b F(t)\,\mathrm{d}t \right) \right) \\
&= \liminf_n \left(n \left(\int_b^{b+\frac{1}{n}} F(t)\,\mathrm{d}t - \int_a^{a+\frac{1}{n}} F(t)\,\mathrm{d}t \right) \right) \\
&\leq \liminf_n \left(n \left(F(b)\frac{1}{n} - F(a)\frac{1}{n} \right) \right) = F(b) - F(a).
\end{aligned}
$$

This proves in turn $F' \in L^1([a,b])$. ■

Remark 4.6: The inequality established may be strict. Recall f the Cantor-Lebesgue function. It is continuous and increasing, so f' exists almost everywhere on $[0,1]$, indeed, for every $x \in [0,1] \setminus C$, $f'(x) = 0$. Then, $f' \equiv 0$ a.e. on $[0,1]$ so $\int_0^1 f'(x)\,\mathrm{d}x = 0$, while $f(1) - f(0) = 1$.

§4.4 Functions of Bounded Variation

↪**Definition 4.3** (Bounded Variation): A function $F : [a,b] \to \mathbb{R}$ is of *bounded variation* on $[a,b]$, denoted by $f \in \mathrm{BV}([a,b])$ if

$$
T_F(a,b) := \sup\left\{ \sum_{k=1}^{N} |F(x_k) - F(x_{k-1})| : N \geq 1, a = x_0 < x_1 < x_2 < \cdots < x_N = b \right\} < \infty.
$$

We call $x_0, \ldots, x_N$ a partition of $[a,b]$, and T_F the *total variation* of F over $[a,b]$.

↪**Proposition 4.6**:

1. For $x \in [a,b]$, set $V_F(x) := T_F(a,x)$. Then, $V_F : [a,b] \to [0,\infty]$ is increasing and for every $a \leq x \leq y \leq b$, $V_F(y) - V_F(x) = T_F((x,y))$.
2. If $F, G \in \mathrm{BV}([a,b])$, then both $F + G, F - G \in \mathrm{BV}([a,b])$ and $T_{F+G}(a,b) \leq T_F(a,b) + T_G(a,b)$.
3. If F is monotonic on $[a,b]$, $F \in \mathrm{BV}([a,b])$.
4. If $f \in L^1(\mathbb{R})$ and $F(x) := \int_a^x f(t)\,\mathrm{d}t$ for $x \in [a,b]$, then $F \in \mathrm{BV}([a,b])$ and $T_F(a,b) \leq \int_a^b |f(t)|\,\mathrm{d}t$.
5. If $F \in \mathrm{BV}([a,b])$, F is bounded on $[a,b]$.
6. If F is continuous on $[a,b]$ and differentiable everywhere on (a,b), and there is some $M > 0$ such that $|F'(x)| \leq M$ for every $x \in (a,b)$, then $F \in \mathrm{BV}([a,b])$.
7. In 6., the boundedness of F' cannot be dropped.

PROOF.

3. For any partition $a = x_0 < \cdots < x_N = b$, we have

$$\sum_{k=1}^{N} |F(x_k) - F(x_{k-1})| = \sum_{k=1}^{N} F(x_k) - F(x_{k-1}) = F(b) - F(a),$$

which was true of any partition so in particular $T_F(a,b) = |F(b) - F(a)|$.

4. For any partition $a = x_0 < \cdots < x_N = b$,

$$\begin{aligned}\sum_{k=1}^{N} |F(x_k) - F(x_{k-1})| &= \sum_{k=1}^{N} \Big| \int_{x_{k-1}}^{x_k} f(t)\,\mathrm{d}t \Big| \\ &\le \sum_{k=1}^{N} \int_{x_{k-1}}^{x_k} |f(t)|\,\mathrm{d}t = \int_a^b |f(t)|\,\mathrm{d}t < \infty.\end{aligned}$$

5. For every $x \in [a,b]$,

$$\begin{aligned}|F(x)| \le |F(x) - F(a)| + |F(a)| &\le T_F(a,x) + |F(a)| \\ &\le T_F(a,b) + |F(a)| < \infty.\end{aligned}$$

6. By the mean value theorem, for every $x < y \in (a,b)$, there is some $z \in (x,y)$ such that $\frac{F(y)-F(x)}{y-x} = F'(z)$. Hence, for every such x, y,

$$|F(y) - F(x)| \le M(y-x).$$

For any partition $a = x_0 < \cdots < x_N = b$, then, $\sum_{k=1}^{N} |F(x_k) - F(x_{k-1})| \le M(b-a)$.

7. For instance, consider $F(x) = \begin{cases} x\sin\left(\frac{1}{x^2}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$. Then F continuous on $[0,1]$ and differentiable on $(0,1)$, but $F \notin \mathrm{BV}([0,1])$.

■

↪**Theorem 4.3**: A function $F : [a,b] \to \mathbb{R}$ is in $\mathrm{BV}([a,b])$ if and only if there exist $H, G : [a,b] \to \mathbb{R}$ increasing such that $F(x) = H(x) - G(x)$ for every $x \in [a,b]$.

Proof. ($\Leftarrow$) Increasing functions are in $\mathrm{BV}([a,b])$, so H, G and thus $H - G \in \mathrm{BV}([a,b])$.

($\Rightarrow$) Assume $F \in \mathrm{BV}([a,b])$. Let $H(x) = V_F(x)$ for $x \in [a,b]$, which is increasing on $[a,b]$. Let $G = H - F$, which we claim is also increasing. For every $x < y \in [a,b]$,

$$\begin{aligned}G(y) - G(x) &= H(y) - H(x) - (F(y) - F(x)) \\ &= T_F(x,y) - (F(y) - F(x)) \\ &\ge T_F(x,y) - |F(y) - F(x)| \ge 0.\end{aligned}$$

■

↪**Theorem 4.4**: If $F : [a,b] \to \mathbb{R}$ is of bounded variation, then F is continuous on $[a,b]$ except at at most countably many points, F' exists almost everywhere on $[a,b]$, and $F' \in L^1([a,b])$.

§4.5 Absolutely Continuous Functions

↪**Definition 4.4** (Absolutely Continuous): A function $F : [a, b] \to \mathbb{R}$ is called *absolutely continuous* on $[a, b]$, denoted $F \in \mathrm{AC}([a, b])$, if for every $\varepsilon > 0$ there is a $\delta > 0$ such that for any disjoint intervals $(a_k, b_k) \subseteq (a, b), k = 1, \ldots, N$ with $\sum_{k=1}^{N}(b_k - a_k) \leq \delta$, it holds that

$$\sum_{k=1}^{N} |F(b_k) - F(a_k)| \leq \varepsilon.$$

Remark 4.7: The $\{(a_k, b_k)\}$'s need not partition (a, b).

↪**Proposition 4.7** (Properties of $\mathrm{AC}([a, b])$):
1. $F \in \mathrm{AC}([a, b]) \Rightarrow F$ is uniformly continuous on $[a, b]$.
2. If $f \in L^1([a, b])$ and $F(x) := \int_a^x f(t)\,\mathrm{d}t$ for $x \in [a, b]$, then $F \in \mathrm{AC}([a, b])$.
3. If $F, G \in \mathrm{AC}([a, b])$ then $F + G, F - G \in \mathrm{AC}([a, b])$.
4. If $F \in \mathrm{AC}([a, b])$, then $F \in \mathrm{BV}([a, b])$.
5. If F is continuous on $[a, b]$ and differentiable everywhere on (a, b) and there is some $M > 0$ such that $|F'(x)| \leq M$ for every $x \in (a, b)$, then $F \in \mathrm{AC}([a, b])$.

Proof.
1. For $\varepsilon > 0$, let δ as in the definition, then for every $x, y \in [a, b], x < y$, if $y - x < \delta$, $|F(y) - F(x)| \leq \varepsilon$ (namely, taking a single interval in the definition of $\mathrm{AC}([a, b])$).
2. Recall that $\forall\, \varepsilon < 0$, there is a constant $M > 0$ such that $\int_{\{f^* > M\}} |f| < \frac{\varepsilon}{2}$. Let $\delta = \frac{\varepsilon}{2M}$. Then, for every $(a_k, b_k) \subseteq (a, b)$ disjoint, $k = 1, \ldots, N$ such that $\sum_{k=1}^{N}(b_k - a_k) \leq \delta$, we have

$$\begin{aligned}
\sum_{k=1}^{N} |F(b_k) - F(a_k)| &\leq \sum_{k=1}^{N} \int_{a_k}^{b_k} |f| \\
&= \sum_{k=1}^{N} \int_{(a_k, b_k) \cap \{f^* > M\}} |f| + \sum_{k=1}^{N} \int_{(a_k, b_k) \cap \{f^* \leq M\}} |f| \\
&\leq \int_{\bigcup_{k=1}^{N}(a_k, b_k) \cap \{f^* > M\}} |f| + M \sum_{k=1}^{N}(b_k - a_k) \\
&\leq \frac{\varepsilon}{2} + M\delta = \varepsilon.
\end{aligned}$$

4. Let $\varepsilon = 1$ and take δ in the definition of AC. Consider a partition of $[a, b]$, $a = t_0 < \cdots < t_L = b$ such that $t_{i+1} - t_i = \frac{b-a}{L}$ and L is such that $\frac{b-a}{L} \leq \delta$. For each $i = 0, \ldots, L$, take any partition of $[t_i, t_{i+1}]$, $t_i = x_0 < \cdots < x_N = t_{i+1}$. Then, (x_k, x_{k+1})'s are disjoint and $\sum_{k=0}^{N-1}(x_{k+1} - x_k) = t_{i+1} - t_i \leq \delta$. So, $\sum_{k=0}^{N-1} |F(x_{k+1}) - F(x_k)| \leq 1$ i.e. $T_F(t_i, t_{i+1}) \leq 1$. Then, $T_F(a, b) \leq \sum_{i=0}^{L-1} T_F(t_i - t_{i+1}) \leq L < \infty$.
5. Use mean value theorem and the similar proof for $\mathrm{BV}([a, b])$.

∎

↪**Theorem 4.5**: If $F \in \mathrm{AC}([a,b])$, then F' exists a.e. on $[a,b]$ and $F' \in L^1([a,b])$.

Proof. $\mathrm{AC}([a,b]) \subseteq \mathrm{BV}([a,b])$, and the same property holds for $\mathrm{BV}([a,b])$. ∎

↪**Theorem 4.6**: Given $F \in \mathrm{AC}([a,b])$, F is constant on $[a,b]$, that is, there is some $c \in \mathbb{R}$ such that $F(x) = c$ for every $x \in [a,b]$ if and only if $F' = 0$ a.e. on $[a,b]$.

Proof. ($\Rightarrow$) If $F \equiv c$, $F' \equiv 0$ on (a,b).

($\Leftarrow$) Assume $F \in \mathrm{AC}([a,b])$ and $F' = 0$ a.e. on $[a,b]$. We want to show that for every $c \in (a,b]$, $F(c) = F(a)$. Fix $c \in (a,b]$. Set $E = \{x \in [a,c] : F'(x) = 0\}$, so $m([a,c] \setminus E) = 0$. Fix $\varepsilon > 0$, let $\delta > 0$ as in the definition of AC, and let

$$\mathcal{J} := \{I = [x, x+h] \subseteq (a,c) : x \in E, h > 0, |F(x+h) - F(x)| \leq \varepsilon h\}.$$

Then, for every $x \in E$, $x \in (a,c)$ and $F'(x) = 0$ so there is some $I = [x, x+h] \in \mathcal{J}$ with $x \in I$ and h arbitrarily small. So in particular, $\mathcal{J}$ a Vitali covering of E. Then, there are disjoint $I_1, ..., I_N \in \mathcal{J}$ such that $m\left(E \setminus \bigcup_{i=1}^N I_i\right) \leq \delta$. Hence, $m\left([a,c] \setminus \bigcup_{i=1}^N I_i\right) \leq \delta$. Denote $I_i = [x_i, x_i + h]$, and relabel so that they are increasing, namely $x_1 < x_1 + h_1 < x_2 < x_2 + h_2 < \cdots < x_N < x_N + h_N$. For every $i = 1, ..., N$, $|F(x_i + h_i) - F(x_i)| \leq \varepsilon h_i$, by construction. So, notice that

$$|F(a) - F(c)| \leq |F(a) - F(x_1)| + \underbrace{\sum_{i=1}^N |F(x_i + h_i) - F(x_i)|}_{\substack{\leq \varepsilon \sum_{i=1}^N h_i \\ \leq \varepsilon(c-a)}} + \sum_{i=1}^{N-1} |F(x_{i+1}) - F(x_i + h_i)| + |F(c) - F(x_N + h_N)|.$$

The remaining intervals to deal with are $(a, x_1), \{(x_i + h_i, x_{i+1})\}, (x_N + h_N, c)$. These are all disjoint, and the union of them equals $(a,c) \setminus \bigcup_{i=1}^N I_i$. Hence, the sum of the lengths of these intervals is bounded by δ. So, $|F(x_1) - F(a)| + \sum_{i=1}^N |F(x_{i+1}) - F(x_i + h_i)| + |F(c) - F(x_N + h_N)| \leq \varepsilon$ by AC. Thus,

$$|F(a) - F(c)| \leq \varepsilon + \varepsilon(c-a) = \varepsilon(c - a + 1),$$

and since ε arbitrarily small, it must be that $F(a) = F(c)$, completing the proof. ∎

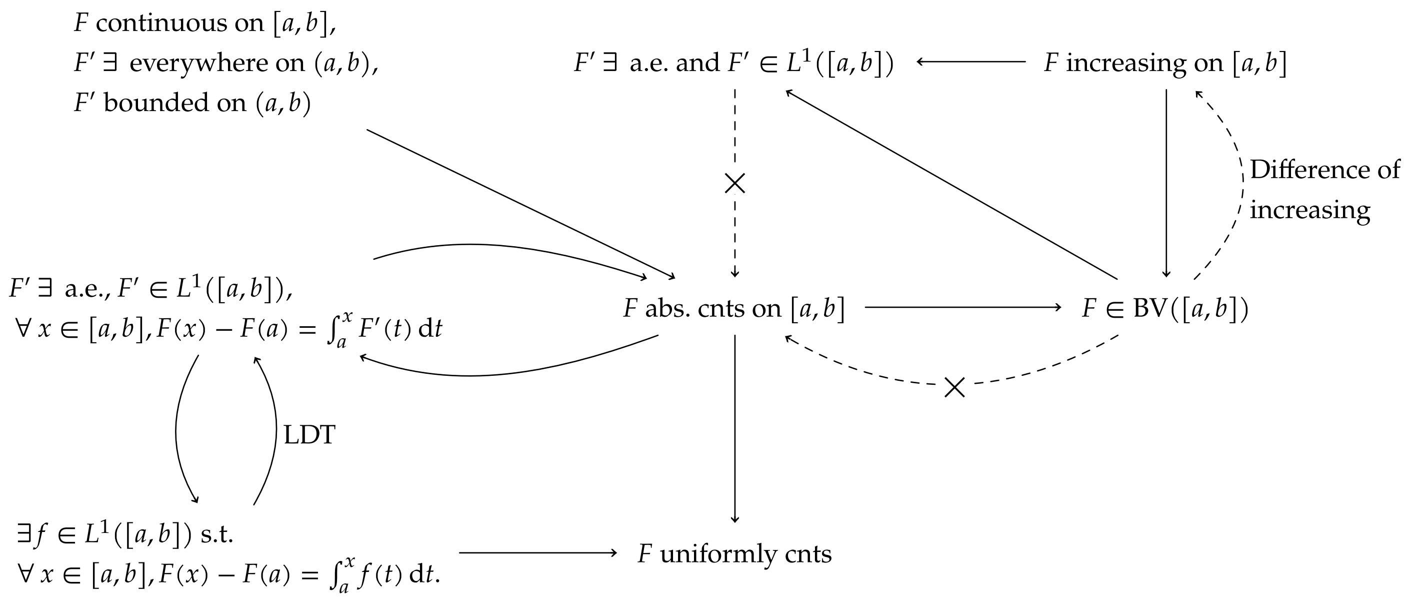
Remark 4.8: The condition $F \in \mathrm{AC}([a,b])$ cannot be dropped. Namely, if F is the Cantor-Lebesgue function on $[0,1]$, then $F' \equiv 0$ a.e. but F is not constant. In particular, F is not $\mathrm{AC}([0,1])$, but is $\mathrm{BV}([0,1])$, being an increasing function.

↪**Theorem 4.7** (Fundamental Theorem of Calculus): If $F \in \mathrm{AC}([a,b])$, then F' exists almost everywhere on $[a,b]$, $F' \in L^1([a,b])$, and for every $x \in [a,b]$,

$$F(x) - F(a) = \int_a^x F'(t)\,\mathrm{d}t.$$

In particular, $F(b) - F(a) = \int_a^b F'(t)\,\mathrm{d}t$.

Proof. Assume $F \in \mathrm{AC}([a,b])$. Define $G(x) := F(a) + \int_a^x F'(t)\,\mathrm{d}t$ for every $x \in [a,b]$. Then, since $F' \in L^1(\mathbb{R})$, $\int_a^x F'(t)\,\mathrm{d}t \in \mathrm{AC}([a,b])$ so $G \in \mathrm{AC}([a,b])$. Moreover, by theorem 4.2, $G' = F'$ almost everywhere on $[a,b]$. Thus, $H := F - G \in \mathrm{AC}([a,b])$ and $H' = F' - G' = 0$ almost everywhere on $[a,b]$ hence $H(x) = H(a) = 0$ for every $x \in [a,b]$. Hence, $F(x) = G(x) = F(a) + \int_a^x F'(t)\,\mathrm{d}t$ for every $x \in [a,b]$. ■



§5 A Glance Towards Probability Theory

Assume μ a probability measure on $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$. Define $F_\mu(x) := \mu((-\infty, x])$ for $x \in \mathbb{R}$, called the distribution function of μ. Then,

1. F_μ increasing.
2. $\lim_{x\to\infty} F_\mu(x) = 1$ and $\lim_{x\to-\infty} F_\mu(x) = 0$.
3. F_μ has at most countably many discontinuities and F_μ is RCLL (right continuous with left-handed limits) i.e. for every $x \in \mathbb{R}$, $F_\mu(x+) = F_\mu(x)$ and $F_\mu(x-)$ exists.
4. For every $x \in \mathbb{R}$, $F_\mu(x) - F_\mu(x-) = \mu(\{x\})$ i.e. F_μ continuous at $x \Leftrightarrow \mu(\{x\}) = 0$.
5. For every $a < b \in \mathbb{R}$, $F_\mu(b) - F_\mu(a) = \mu((a,b])$, $F_\mu(b-) - F_\mu(a) = \mu((a,b))$, $F_\mu(b) - F_\mu(a-) = \mu([a,b])$, $F_\mu(b-) - F_\mu(a-) = \mu([a,b))$.
6. F_μ uniquely determines μ, i.e. if μ, ν are both probability measures on $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$ and $F_\mu = F_\nu$ then $\mu = \nu$.
7. Any $F : \mathbb{R} \to [0,1]$ satisfying 1., 2., and 3. is the distribution function of some probability measure μ on $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$.

↪**Definition 5.1**: Let μ a probability measure on $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$ with df F_μ. We say μ is absolutely continuous with respect to m and write $\mu \ll m$ if F_μ is absolutely continuous on $\mathbb{R}$. I.e., there is some $f \in L^1(\mathbb{R})$ such that for every $a \in \mathbb{R}, x > a$, $F_\mu(x) - F_\mu(a) = \int_a^x f(t)\,\mathrm{d}t$. Such an f is called the "probability density" of μ.

↪**Proposition 5.1**: If μ absolutely continuous wrt m with density f then for every $B \in \mathfrak{B}_{\mathbb{R}}$, $\mu(B) = \int_B f$.

Proof. Let $\tilde{\mu}$ on $\mathfrak{B}_{\mathbb{R}}$ by $\tilde{\mu}(B) := \int_B f$ for every $B \in \mathfrak{B}_{\mathbb{R}}$. One can verify $\tilde{\mu}$ a probability measure on $\mathbb{R}$. If $F_{\tilde{\mu}}$ the distribution function of $\tilde{\mu}$, then for every $x \in \mathbb{R}$, $F_{\tilde{\mu}}(x) = F(x) \Rightarrow \mu = \tilde{\mu}$. ■

↪**Corollary 5.1**: Assume $\mu \ll m$ with density f. Then if $g : \mathbb{R} \to \mathbb{R}$ Borel-measurable, then $\int_{\mathbb{R}} |g|\,\mathrm{d}\mu = \int_{\mathbb{R}} |g(x)| f(x)\,\mathrm{d}x$. In particular, $g \in L^1(\mu) \Leftrightarrow g \cdot f \in L^1(\mathbb{R})$.

Remark 5.1: Any equivalent description of $\mu \ll m$ is that for every $A \in \mathfrak{B}_{\mathbb{R}}$, if $m(A) = 0, \mu(A) = 0$. More generall:

↪**Theorem 5.1** (Radon-Nikodym Theorem): Let μ, ν be two σ-finite measures on $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$. If $\forall\, A \in \mathfrak{B}_{\mathbb{R}}\ \nu(A) = 0 \Rightarrow \mu(A) = 0$ then $\mu \ll \nu$ and so there is some $f \in L^1(\nu)$ such that for every $B \in \mathfrak{B}_{\mathbb{R}}$, $\mu(B) = \int_B f\,\mathrm{d}\nu$. We call such an f the Radon-Nikodym derivative of μ with respect to ν, denoted by $f = \frac{\mathrm{d}\mu}{\mathrm{d}\nu}$.

Remark 5.2: On the other hand, if μ is such that $\mu(B) = \int_B f \, d\nu$ for some $f \in L^1(\nu)$ and for every $B \in \mathfrak{B}_{\mathbb{R}}$, then ν-null $\Rightarrow$ μ-null.

↪**Theorem 5.2** (Lebesgue Decomposition Theorem): Given μ any probability measure on $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$, μ admits a *unique* decomposition $\mu = \mu_a + \mu_s$ such that

1. μ_a a finite measure on $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$ and $\mu_a \ll m$.
2. μ_s a finite measure on $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$ such that "$\mu_s \perp m$", that is μ_s "singular" to m i.e. there exists $E \in \mathfrak{B}_{\mathbb{R}}$ such that $m(E) = 0$ but $\mu_s(E^c) = 0$.

Proof. Set $\lambda = \mu + m$. Then, λ a σ-finite measure on $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$ and for every $A \in \mathfrak{B}_{\mathbb{R}}$, if $\lambda(A) = 0$, then $\mu(A) = m(A) = 0$. By R-N Thm, there is some $f, g \in L^1(\lambda)$ such that $\mu(B) = \int_B f \, d\lambda, m(B) = \int_B g \, d\lambda$ for every $B \in \mathfrak{B}_{\mathbb{R}}$.

Set $E := \{x \in \mathbb{R} : g(x) = 0\}$. Then, $m(E) = 0$. Define μ_a, μ_s by for every $B \in \mathfrak{B}_{\mathbb{R}}$ by $\mu_a(B) = \mu(B \cap E^c), \mu_s(B) = \mu(B \cap E)$.

Then, $\mu = \mu_a + \mu_s$ and $\mu_s(E^c) = \mu(E^c \cap E) = 0$.

We need to show $\mu_a \ll m$. Assume $A \in \mathfrak{B}_{\mathbb{R}}$ is such that $m(A) = 0$. Then,

$$0 = \int_A g \, d\lambda = \cancel{\int_{A \cap E} g \, d\lambda} + \int_{A \cap E^c} g \, d\lambda \Rightarrow \lambda(A \cap E^c) = 0,$$

so $\mu(A \cap E^c) = \mu_a(A) = 0$. ∎

⊛ **Example 5.1**: Let $F : \mathbb{R} \to [0, 1]$ be 1 for $x \geq 1$, 0 for $x \leq 0$, and the Cantor Lebesgue function on $[0, 1]$. Then, F is a distribution function of a (unique) probability measure μ. In fact, $\mu \perp m$. For instance, if C is the Cantor set, $m(C) = 0$ and $\mu(\mathbb{R} \setminus C) = 0$.