MATH 501 ANALYSIS

BURAK KAYA

Abstract. These are the lecture notes I used for a 14-week introductory graduate-level analysis class that I taught at the Department of Mathematics of Middle East Technical University during Fall 2019.

Contents

In order to determine the course content and prepare these lecture notes, I mainly used the classic textbook of Folland [Fol99] which I also listed as the textbook for the course. The aim of the course was to **fully or partly** cover Sections §1.1, 1.2, 1.3, 1.4, 1.5, 2.1, 2.2, 2.3, 2.4, 2.5, 3.1 and 3.2 of Folland's book [Fol99]. However, I also covered some additional material which I think is of importance.

Besides [Fol99], I also listed the supplementary resources [Coh93] and [Bog07], whose e-book versions can be downloaded from this link and this link respectively. Indeed, I occasionally followed [Coh93] to cover certain topics. During the lectures, besides what is included here, many additional instructional examples were considered, some of which unfortunately could not make it into these lectures notes.

0. Prelude

0.1. **Why is Riemann integral not sufficient?** One is usually exposed to integration theory for the first time via Riemann integrals, which, for most purposes in practice, are sufficient.

On the other hand, if one does more theoretical (and serious) mathematics, then one realizes that the Riemann integral lacks various "nice" properties, which one usually wishes to have. To illustrate such this, let us first recall the definition of Riemann integrability.

Given a compact interval $[a,b]$ and a bounded function $f:[a,b]\to\mathbb{R}$, we say that f is Riemann integrable over $[a,b]$ if and only if

$$\inf\{U(f,P) : P \text{ is a partition of } [a,b]\} = \sup\{L(f,P) : P \text{ is a partition of } [a,b]\}$$

where, for a partition $P=\{t_k : 0\le k\le n\}$ of $[a,b]$ with $a=t_0<t_1<\cdots<t_n=b$, we write $\Delta t_k=(t_k-t_{k-1})$ and

$$U(f,P)=\sum_{k=1}^{n}\left(\sup_{x\in[t_{k-1},t_k]} f(x)\right)\cdot\Delta t_k \quad\text{and}\quad L(f,P)=\sum_{k=1}^{n}\left(\inf_{x\in[t_{k-1},t_k]} f(x)\right)\cdot\Delta t_k$$

It is a standard calculus fact that any continuous function is Riemann integrable over a compact interval. Unfortunately, working only with continuous functions is too restrictive if we are to do more than computing areas of plane regions.

To see an example of a non-Riemann integrable function, set $A=\mathbb{Q}\cap[0,1]$ and consider the characteristic function $\chi_A:[0,1]\to\mathbb{R}$ defined by

$$\chi_A(x)=\begin{cases}1 & \text{if } x\in A\\ 0 & \text{if } x\notin A\end{cases}$$

Since rational and irrational numbers are both dense sets in $\mathbb{R}$, a moment's thought reveals that $U(\chi_A,P)=1$ and $L(\chi_A,P)=0$ for any partition P of $[0,1]$. It follows that χ_A is not Riemann integrable over $[0,1]$.

Let us now enumerate A, say, $A=\{a_i : i\in\mathbb{N}\}$ and set $A_n=\{a_i : 0\le i\le n\}$. On the one hand, the characteristic functions $\chi_{A_n}:[0,1]\to\mathbb{R}$ *are* Riemann integrable over $[0,1]$ and indeed

$$\int_0^1 \chi_{A_n}(x)dx = 0$$

On the other hand, for every $x \in [0,1]$, we have that $\lim_{n\to\infty} \chi_{A_n}(x) = \chi_A(x)$ but

$$\lim_{n\to\infty} \int_0^1 \chi_{A_n}(x)dx \neq \int_0^1 \lim_{n\to\infty} \chi_{A_n}(x)dx = \int_0^1 \chi_A(x)dx$$

as the right-hand side does not even exist. In other words, one cannot simply interchange limit and the Riemann integral, even if the functions whose limit is to be taken converge in a "nice" (for example, *monotone*) way.

We wish to build a better theory of integration where one can pull off such tricks. What should be our starting point? Since the last obstacle we just had arose from integrals of characteristic functions, maybe we should take care of these first. Let $A \subseteq \mathbb{R}$ be an arbitrary set. What should $\int \chi_A$ be?

Let us take a look at what we already have. In the case that A is an interval, we have that the Riemann integral $\int \chi_A(x)dx$ is the *length* of A. We also want our to-be-defined-later integral to extend the Riemann integral. This means that we should come up with an integration method such that $\int \chi_A$ is the "length" of A for *any* subset of $\mathbb{R}$. How can one "measure the lengths" of arbitrary subsets of $\mathbb{R}$?

0.2. **To be measurable or not to be measurable, that is the question.** We wish to have a method of measuring the lengths (areas, volumes etc.) of all subsets of $\mathbb{R}^n$. Such a method would be a function $\mu : \mathcal{P}(\mathbb{R}^n) \to [0,+\infty]$. Let us focus on the case $n = 1$ for simplicity.

What are our expectations from such a function? For the purposes of doing calculus, we want $\mu : \mathcal{P}(\mathbb{R}) \to [0,+\infty]$ to be such that

a. If $\{A_i : i \in \mathbb{N}\}$ are disjoint subsets of $\mathbb{R}$, then $\mu\left(\bigcup_{i=0}^{\infty} A_i\right) = \sum_{i=0}^{\infty} \mu(A_i)$.
b. $\mu(A) = \mu(x + A)$ for any $x \in \mathbb{R}$ and $A \subseteq R$.
c. $\mu((a,b)) = b - a$ for any $a, b \in \mathbb{R}$ with $a \leq b$.

The property (a) is called *σ-additivity.* If one replaces the countable collection of sets in (a) by a finite collection of sets, then the corresponding property is called *finite additivity.* The property (b) is called *translation-invariance.*

In other words, we wish to have a map $\mu : \mathcal{P}(\mathbb{R}) \to [0,+\infty]$ which is σ-additive, translation-invariant function and assigns closed intervals their lengths. Unfortunately, one cannot always get what one wants.

Theorem 1 (Vitali)**.** *There does not exist a σ-additive translation-invariant map $\mu : \mathcal{P}(\mathbb{R}) \to [0,+\infty]$ such that $0 < \mu([0,1)) < +\infty$.*

Proof. Assume that there exists such a map μ. Consider the action of $\mathbb{Z}$ on $[0,1)$ given by $n \cdot x = (n\sqrt{2} + x \pmod 1)$. Note that this action is free, that is, $m \cdot x = n \cdot x$ implies that $m = n$. This is true because if $m \cdot x = n \cdot x$, then $(m-n)\sqrt{2} \equiv 0 \pmod 1$, which, together with the irrationality of $\sqrt{2}$, implies that $m = n$.

Using the axiom of choice, one can show that there exists a transversal set $T \subseteq [0,1)$ for the orbit equivalence relation of the action $\mathbb{Z} \curvearrowright [0,1)$. Since the action is free, the translates of the transversal of its orbit equivalence relation are disjoint, that is, $n \cdot T \cap m \cdot T = \emptyset$ whenever $m \neq n$. It follows that

$$[0,1) = \bigsqcup_{n \in \mathbb{Z}} n \cdot T$$

Note that, since μ is translation-invariant and σ-additive, we have $\mu(T) = \mu(n \cdot T)$ for any $n \in \mathbb{Z}$. It then follows from the properties of ν that

$$+\infty > \mu([0,1)) = \mu\left(\bigsqcup_{n \in \mathbb{Z}} n \cdot T\right) = \sum_{n \in \mathbb{Z}} \mu(n \cdot T) = \sum_{n \in \mathbb{Z}} \mu(T) > 0$$

which is a contradiction as there can be no such number $\mu(T)$. Therefore, there exists no such function μ. □

The idea in this proof is due to Vitali and is usually carried out by the action of $\mathbb{Q}$ on $\mathbb{R}$ via left-translation. The corresponding transversal sets are called *Vitali sets* and have to be "non-measurable" just like the set T in the proof above.

Now that we know there can be no such function μ, the next step will be to find the guilty. The properties (b) and (c) are indispensable if one is to have a geometrically meaningful theory of integration. Thus we may try to relax the property (a). In order to carry out basic calculus, say, to split an integral as $\int_a^b = \int_a^c + \int_c^b$, the best one can demand is to relax σ-additivity to finite additivity.

It is a remarkable fact that the Hahn-Banach theorem implies that, for $n = 1$ (respectively, $n = 2$), there **does** exist a *finitely additive* isometry-invariant function $\mu : \mathcal{P}(\mathbb{R}^n) \to [0, +\infty]$ which assigns closed intervals (respectively, closed rectangles) their lengths (respectively, their areas.)

A more remarkable fact is that this cannot be done in higher dimension. The Banach-Tarski theorem states that a unit closed ball in $\mathbb{R}^3$ can be partitioned into five pieces so that one can obtain *two* unit closed balls by applying translations and rotations of $\mathbb{R}^3$ to these pieces. Consequently, there can be no such finitely additive measure μ for $n = 3$. (See [Wag93] for an excellent monograph on this theorem and related topics.) Therefore, that we cannot measure all subsets of $\mathbb{R}^n$ in a meaningful way is not due to strong additivity assumptions. What is the problem then?

Maybe we should not demand to measure *all* subsets of $\mathbb{R}^n$ but only demand to measure a reasonably rich collection of subsets that actually arise in mathematics. This brings us to the notions of an algebra and σ-algebra, which will be the collections of sets that we are going to measure.

1. Algebras

1.1. **Algebras and σ-algebras.** Let $\mathbf{X}$ be a non-empty set. An *algebra* on $\mathbf{X}$ is a collection $\mathcal{A} \subseteq \mathcal{P}(\mathbf{X})$ such that

- $\emptyset \in \mathcal{A}$.
- $\mathcal{A}$ is closed under complements, that is, if $A \in \mathcal{A}$, then $A^c = \mathbf{X} - A \in \mathcal{A}$.
- $\mathcal{A}$ is closed under finite unions, that is, if $A_1, \ldots, A_n \in \mathcal{A}$, then $\bigcup_{i=1}^n A_i \in \mathcal{A}$.

It is easily seen that an algebra on $\mathbf{X}$ also contains $\mathbf{X}$, is closed under finite intersections as $\bigcap_{i=1}^n A_i = (\bigcup_{i=1}^n A_i^c)^c$ and is closed under set differences as $A-B = A \cap B^c$. A *σ-algebra* on $\mathbf{X}$ is an algebra which is also closed under countable unions, that is, $\mathcal{A} \subseteq \mathcal{P}(\mathbf{X})$ is a σ-algebra if

- $\emptyset \in \mathcal{A}$,
- $A \in \mathcal{A}$ implies $A^c = \mathbf{X} - A \in \mathcal{A}$, and
- $A_1, A_2, \cdots \in \mathcal{A}$ implies $\bigcup_{i=1}^\infty A_i \in \mathcal{A}$.

Algebras (respectively, σ-algebras) will later serve as the domains of "measures". This is why we expect these collections to be closed under finite and countable unions respectively. More precisely, the domain of a finitely (respectively, σ-)additive measure should be closed under finite (respectively, countable) unions.

Suppose that $\mathcal{A}$ is an algebra which is closed under *disjoint* countable unions. We claim that $\mathcal{A}$ is indeed a σ-algebra. Let $A_1, A_2, \cdots \in \mathcal{A}$. Set $B_1 = A_1$ and $B_k = A_k - \bigcup_{n=1}^{k-1} A_n$ for $k \geq 2$. Since $\mathcal{A}$ is an algebra, we have that $B_k \in \mathcal{A}$ for all $k \geq 1$. Moreover, $\bigcup_{k=1}^\infty A_k = \bigcup_{k=1}^\infty B_k \in \mathcal{A}$. Therefore $\mathcal{A}$ is closed under countable unions and hence, is a σ-algebra.

Here are some examples of σ-algebras on a set $\mathbf{X}$.

- $\mathcal{A} = \mathcal{P}(\mathbf{X})$.
- $\mathcal{A} = \{\emptyset, \mathbf{X}\}$.
- $\mathcal{A} = \{\emptyset, A, B, \mathbf{X}\}$ where $\{A, B\}$ is a partition of $\mathbf{X}$.
- $\mathcal{A} = \{A \subseteq \mathbf{X} : A \text{ or } A^c \text{ is uncountable}\}$.

Next will be introduced the notion of a *generating set* of a σ-algebra. In order to define this, we need the following proposition.

Proposition 1. *Let $\mathbf{X}$ be a non-empty set and let $\{\mathcal{A}_i\}_{i\in I}$ be σ-algebras on $\mathbf{X}$. Then $\bigcap_{i\in I} \mathcal{A}_i$ is a σ-algebra on $\mathbf{X}$.*

Proof. Clearly $\emptyset \in \bigcap_{i\in I} \mathcal{A}_i$ since $\emptyset \in \mathcal{A}_i$ for every $i \in I$. Let $A, A_1, A_2, \cdots \in \bigcap_{i\in I} \mathcal{A}_i$. Then, by definition, $A, A_1, A_2, \cdots \in \mathcal{A}_i$ for every $i \in I$. Since $\mathcal{A}_i$ is a σ-algebra for every $i \in I$, we have that $A^c \in \mathcal{A}_i$ and $\bigcup_{n=1}^\infty A_n \in \mathcal{A}_i$ for every $i \in I$. Thus $A^c \in \bigcap_{i\in I} \mathcal{A}_i$ and $\bigcup_{n=1}^\infty A_n \in \bigcap_{i\in I} \mathcal{A}_i$. □

Given $\mathcal{E} \subseteq \mathcal{P}(\mathbf{X})$, define

$$\mathcal{M}(\mathcal{E}) = \bigcap \{\mathcal{A} \subseteq \mathcal{P}(\mathbf{X}) : \ \mathcal{E} \subseteq \mathcal{A} \ \text{ and } \mathcal{A} \text{ is a } \sigma\text{-algebra}\}$$

Since there always exists a σ-algebra containing $\mathcal{E}$, namely, $\mathcal{P}(\mathbf{X})$, the collection on the right-hand side is non-empty and hence its intersection is defined. The collection $\mathcal{M}(\mathcal{E})$ is a σ-algebra by Proposition 1. The σ-algebra $\mathcal{M}(\mathcal{E})$ is the smallest σ-algebra containing $\mathcal{E}$ and is called the σ-algebra *generated by* $\mathcal{E}$.

Observe that, for any $\mathcal{E}, \mathcal{F} \subseteq \mathcal{P}(\mathbf{X})$, if $\mathcal{E} \subseteq \mathcal{M}(\mathcal{F})$, then $\mathcal{M}(\mathcal{E}) \subseteq \mathcal{M}(\mathcal{F})$ since $\mathcal{M}(\mathcal{F})$ is a σ-algebra. This basic observation provides us a way to show that σ-algebras generated by two distinct sets are the same: If every element in one of the sets is in the σ-algebra generated by the other set, then these sets generate the same σ-algebra.

We shall now introduce a class of σ-algebras that are of most importance for this course. Let $\mathbf{X}$ be a topological space. The *Borel σ-algebra* of $\mathbf{X}$, shown by $\mathcal{B}(\mathbf{X})$, is the σ-algebra generated by the open sets of $\mathbf{X}$. The elements of $\mathcal{B}(\mathbf{X})$ are called the *Borel sets* of $\mathbf{X}$.

One may think that this definition, although pretty basic, is somewhat implicit in the sense that it does not tell one how to obtain Borel sets. How does the class of Borel sets of a topological space look like?

1.2. **The Borel hierarchy.** In this subsection, we shall stratify the Borel σ-algebra of a metrizable topological space. Due to the nature of our construction, the reader is expected to be familiar with ordinals and transfinite recursion. Those who are not well-read in set theory may skip this subsection for it will not play a crucial role in the remaining of these notes.

Let $\mathbf{X}$ be a metrizable topological space. For every countable ordinal $1 \leq \alpha < \omega_1$, we define the following collections of subsets of $\mathbf{X}$ by transfinite recursion:

- $\mathbf{\Sigma}^0_1 = \{U \subseteq \mathbf{X} : U \text{ is open}\}$,
- $\mathbf{\Pi}^0_\alpha = \{S^c : S \in \mathbf{\Sigma}^0_\alpha\}$ and $\mathbf{\Delta}^0_\alpha = \mathbf{\Sigma}^0_\alpha \cap \mathbf{\Pi}^0_\alpha$ for every $1 \leq \alpha < \omega_1$, and
- $\mathbf{\Sigma}^0_\alpha = \{\bigcup_{n\in\mathbb{N}} A_n : A_n \in \mathbf{\Pi}^0_{\gamma_n}, 1 \leq \gamma_n < \alpha, n \in \mathbb{N}\}$ for every $1 < \alpha < \omega_1$.

The sets in $\mathbf{\Delta}^0_1$, $\mathbf{\Sigma}^0_1$, $\mathbf{\Pi}^0_1$, $\mathbf{\Sigma}^0_2$ and $\mathbf{\Pi}^0_2$ are classically called the clopen, open, closed, F_σ and G_δ subsets of X respectively.

Lemma 1. $\mathbf{\Sigma}^0_\alpha, \mathbf{\Pi}^0_\alpha \subseteq \mathbf{\Delta}^0_{\alpha+1}$ *for every* $1 \leq \alpha < \omega_1$.

Proof. We shall prove this by transfinite induction on $\alpha \geq 1$. That $\mathbf{\Pi}^0_1 \subseteq \mathbf{\Sigma}^0_2$ (and hence $\mathbf{\Sigma}^0_1 \subseteq \mathbf{\Pi}^0_2$) follows from the definition and that $\mathbf{\Pi}^0_1 \subseteq \mathbf{\Pi}^0_2$ (and hence $\mathbf{\Sigma}^0_1 \subseteq \mathbf{\Sigma}^0_2$) follows from the metrizability of $\mathbf{X}$. Thus the claim holds for $\alpha = 1$. Now, let $1 < \alpha < \omega_1$ and assume that the claim holds for all ordinals $1 \leq \theta < \alpha$. We wish to show that the claim also holds for α.

Let $A \in \mathbf{\Sigma}^0_\alpha$. Then $A = \bigcup_{n\in\mathbb{N}} A_n$ for some $A_n \in \mathbf{\Pi}^0_{\gamma_n}$ and some $1 \leq \gamma_n < \alpha$. By inductive assumption, $A_n \in \mathbf{\Pi}^0_{\gamma_n} \subseteq \mathbf{\Delta}^0_{\gamma_n+1} \subseteq \mathbf{\Pi}^0_{\gamma_n+1}$ for all $n \in \mathbb{N}$. Thus, we have that $A = \bigcup_{n\in\mathbb{N}} A_n \in \mathbf{\Sigma}^0_{\alpha+1}$. This shows that $\mathbf{\Sigma}^0_\alpha \subseteq \mathbf{\Sigma}^0_{\alpha+1}$ and hence $\mathbf{\Pi}^0_\alpha \subseteq \mathbf{\Pi}^0_{\alpha+1}$. By definition, we already have that $\mathbf{\Pi}^0_\alpha \subseteq \mathbf{\Sigma}^0_{\alpha+1}$ and hence $\mathbf{\Sigma}^0_\alpha \subseteq \mathbf{\Pi}^0_{\alpha+1}$. These together imply that $\mathbf{\Sigma}^0_\alpha, \mathbf{\Pi}^0_\alpha \subseteq \mathbf{\Delta}^0_{\alpha+1}$ which completes the transfinite induction. $\square$

As a consequence of Lemma 1, these point classes can be pictured as follows, where every class in the diagram is a subset of the classes on right of it.

$$\begin{array}{cccccccccccc} & \mathbf{\Sigma}^0_1 & & \mathbf{\Sigma}^0_2 & & \mathbf{\Sigma}^0_3 & & \cdots & & \mathbf{\Sigma}^0_\alpha & & \\ \mathbf{\Delta}^0_1 & & \mathbf{\Delta}^0_2 & & \mathbf{\Delta}^0_3 & & \cdots & & \mathbf{\Delta}^0_\alpha & & & \alpha < \omega_1 \\ & \mathbf{\Pi}^0_1 & & \mathbf{\Pi}^0_2 & & \mathbf{\Pi}^0_3 & & \cdots & & \mathbf{\Pi}^0_\alpha & & \end{array}$$

We shall now show that the Borel σ-algebra of $\mathbf{X}$ is the union of these point classes and these point classes together form what is known as the *Borel hierarchy* of $\mathbf{X}$.

Theorem 2. $\mathcal{B}(\mathbf{X}) = \bigcup\limits_{1 \leq \alpha < \omega_1} \mathbf{\Sigma}^0_\alpha = \bigcup\limits_{1 \leq \alpha < \omega_1} \mathbf{\Pi}^0_\alpha = \bigcup\limits_{1 \leq \alpha < \omega_1} \mathbf{\Delta}^0_\alpha.$

Proof. We will first prove by transfinite induction that $\mathbf{\Sigma}^0_\alpha \subseteq \mathcal{B}(\mathbf{X})$ for $1 \leq \alpha < \omega_1$. The claim is clearly true for $\alpha = 1$. Now, let $1 < \alpha < \omega_1$ and assume that the claim holds for all ordinals $1 \leq \theta < \alpha$. Let $A \in \mathbf{\Sigma}^0_\alpha$. Then $A = \bigcup_{n \in \mathbb{N}} A_n$ for some $A_n \in \mathbf{\Pi}^0_{\gamma_n}$ and some $1 \leq \gamma_n < \alpha$. By inductive assumption, $A_n^c \in \mathbf{\Sigma}^0_{\gamma_n} \subseteq \mathcal{B}(\mathbf{X})$ for all $n \in \mathbb{N}$ and, $\mathcal{B}(\mathbf{X})$ being a σ-algebra, we have that $A \in \mathcal{B}(\mathbf{X})$. Hence the claim holds for α. Thus, by transfinite induction, $\bigcup_{1 \leq \alpha < \omega_1} \mathbf{\Sigma}^0_\alpha \subseteq \mathcal{B}(\mathbf{X})$.

$\bigcup_{1 \leq \alpha < \omega_1} \mathbf{\Sigma}^0_\alpha$ is clearly closed under complementation and contains $\emptyset$. To show that it is a σ-algebra, let $A_1, A_2, \dots \in \bigcup_{1 \leq \alpha < \omega_1} \mathbf{\Sigma}^0_\alpha$. Then, for every $k \in \mathbb{N}^+$, there exists $1 \leq \alpha_k < \omega_1$ such that $A_k \in \mathbf{\Sigma}^0_{\alpha_k}$. Since ω_1 is a regular cardinal, we have that $\sup\{\alpha_k : k \in \mathbb{N}^+\} = \theta < \omega_1$. It follows that $A_k \in \mathbf{\Sigma}^0_\theta$ for all $k \in \mathbb{N}^+$ and hence $A = \bigcup_{k=1}^\infty A_k \in \mathbf{\Sigma}^0_\theta$ as $\mathbf{\Sigma}^0_\alpha$ are closed under countable unions. Thus $\bigcup_{1 \leq \alpha < \omega_1} \mathbf{\Sigma}^0_\alpha$ is a σ-algebra. Since it contains the open sets of $\mathbf{X}$ by definition, we have that $\mathcal{B}(\mathbf{X}) \subseteq \bigcup_{1 \leq \alpha < \omega_1} \mathbf{\Sigma}^0_\alpha$, which completes the proof of the first equality. The other equalities easily follow from the first one by Lemma 1. □

It is a well known and non-trivial fact that, for an uncountable Polish space, i.e. a separable completely metrizable topological space, the Borel hierarchy does not "collapse" in the sense that $\mathbf{\Sigma}^0_\alpha \neq \mathbf{\Sigma}^0_{\alpha+1}$. We refer the curious reader to [Kec95] for a general background in descriptive set theory, which, in the broadest sense, is the study of the structure of Borel (and projective) sets of Polish spaces.

An immediate consequence of Theorem 2 is that the cardinality of $\mathcal{B}(\mathbb{R})$ is the same as that of $\mathbb{R}$. Let us denote the cardinal $|\mathbb{R}|$ by $\mathfrak{c}$ and call it the *continuum*.

Theorem 3. $|\mathcal{B}(\mathbb{R})| = \mathfrak{c}$.

Proof. We shall prove by transfinite induction that $|\mathbf{\Sigma}^0_\alpha| = \mathfrak{c}$ for every $1 \leq \alpha < \omega_1$. We will first prove this claim for $\alpha = 1$. The map $x \mapsto (0, x)$ from $\mathbb{R}$ to $\mathbf{\Sigma}^0_1$ is clearly injective and hence $\mathfrak{c} \leq |\mathbf{\Sigma}^0_1|$. To show the converse inequality, let $(\mathbf{q}_i)_{i \in \mathbb{N}}$ be an enumeration of $\mathbb{Q}$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, for every open set $O \subseteq \mathbb{R}$ and every $x \in O$, one can choose $q_{x,O}, r_{x,O} \in \mathbb{Q}$ such that $x \in B(q_{x,O}, r_{x,O}) \subseteq O$. Then we have that $O = \bigcup_{x \in O} B(q_{x,O}, r_{x,O})$. Since there are countably many pairs of the form $(q_{x,O}, r_{x,O})$, one can replace this union by a countable union of the form $O = \bigcup_{i\in\mathbb{N}} B(\mathbf{q}_{k_i^O}, \mathbf{q}_{l_i^O})$ for some sequences $(k_i^O)_{i\in\mathbb{N}}, (l_i^O)_{i\in\mathbb{N}} \in \mathbb{N}^{\mathbb{N}}$. For each open $O \subseteq \mathbb{R}$, choose such sequences $(\mathbf{q}_{k_i^O})_{i\in\mathbb{N}}, (\mathbf{q}_{l_i^O})_{i\in\mathbb{N}}$ in $\mathbb{Q}^{\mathbb{N}}$ and consider the map $O \mapsto ((\mathbf{q}_{k_i^O})_{i\in\mathbb{N}}, (\mathbf{q}_{l_i^O})_{i\in\mathbb{N}})$ from $\mathbf{\Sigma}_1^0$ to $\mathbb{Q}^{\mathbb{N}} \times \mathbb{Q}^{\mathbb{N}}$. This map is clearly injective and hence $|\mathbf{\Sigma}_1^0| \leq |\mathbb{Q}^{\mathbb{N}} \times \mathbb{Q}^{\mathbb{N}}| = \mathfrak{c}$. This completes the proof that $|\mathbf{\Sigma}_1^0| = \mathfrak{c}$.

Now let $1 < \alpha < \omega_1$ and assume that the claim holds for all $1 \leq \theta < \alpha$. Since $\mathbf{\Sigma}_1^0 \subseteq \mathbf{\Sigma}_\alpha^0$, we have that $\mathfrak{c} = |\mathbf{\Sigma}_1^0| \leq |\mathbf{\Sigma}_\alpha^0|$. By definition, for every $A \in \mathbf{\Sigma}_\alpha^0$, we have that $A = \bigcup_{n\in\mathbb{N}} A_n$ for some $1 \leq \gamma_n < \alpha$ and some $A_n \in \mathbf{\Pi}_{\gamma_n}^0$. For every $A \in \mathbf{\Sigma}_\alpha^0$, choose such A_n and consider the map

$$A \mapsto (A_n)_{n\in\mathbb{N}} \text{ from } \mathbf{\Sigma}_\alpha^0 \text{ to } \left(\bigcup_{1\leq\theta<\alpha} \mathbf{\Pi}_\theta^0\right)^{\mathbb{N}}$$

This map is clearly injective. Moreover, by inductive assumption, we have that $|\mathbf{\Pi}_\theta^0| = |\mathbf{\Sigma}_\theta^0| = \mathfrak{c}$ for every $1 \leq \theta < \alpha$ and hence the cardinality of $\left(\bigcup_{1\leq\theta<\alpha} \mathbf{\Pi}_\theta^0\right)^{\mathbb{N}}$ is less than or equal to

$$(\mathfrak{c} \cdot |\alpha|)^{\aleph_0} = \max\{\mathfrak{c}, \aleph_0\}^{\aleph_0} = (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0\cdot\aleph_0} = 2^{\aleph_0} = \mathfrak{c}$$

It follows that $|\mathbf{\Sigma}_\alpha^0| \leq \mathfrak{c}$, which completes the inductive step. Thus, by transfinite induction, we have that $|\mathbf{\Sigma}_\alpha^0| = \mathfrak{c}$ for every $1 \leq \alpha < \omega_1$. Now, Theorem 2 implies that

$$\mathfrak{c} \leq |\mathcal{B}(\mathbb{R})| \leq \mathfrak{c} \cdot \omega_1 = \max\{\omega_1, \mathfrak{c}\} = \mathfrak{c}$$

□

1.3. **Generating the Borel sets of $\mathbb{R}$.** We shall now provide some generating sets for the Borel σ-algebra of $\mathbb{R}$ endowed with its usual Euclidean topology.

Proposition 2. *The Borel σ-algebra of $\mathbb{R}$ is generated by the following collections.*

- $\mathcal{E}_1 = \{(a, b) : a, b \in \mathbb{R}\}$
- $\mathcal{E}_2 = \{(a, b) : a, b \in \mathbb{Q}\}$
- $\mathcal{E}_3 = \{[a, b] : a, b \in \mathbb{R}\}$
- $\mathcal{E}_4 = \{[a, b) : a, b \in \mathbb{R}\}$
- $\mathcal{E}_5 = \{(a, b] : a, b \in \mathbb{R}\}$
- $\mathcal{E}_6 = \{(a, \infty) : a \in \mathbb{R}\}$
- $\mathcal{E}_7 = \{(-\infty, a) : a \in \mathbb{R}\}$
- $\mathcal{E}_6 = \{[a, \infty) : a \in \mathbb{R}\}$
- $\mathcal{E}_7 = \{(-\infty, a] : a \in \mathbb{R}\}$

Proof. Let $\mathcal{E}$ be the collection of open sets of $\mathbb{R}$. Then $\mathcal{E}_2 \subseteq \mathcal{E} \subseteq \mathcal{M}(\mathcal{E})$ and consequently $\mathcal{M}(\mathcal{E}_2) \subseteq \mathcal{M}(\mathcal{E}) = \mathcal{B}(\mathbb{R})$. Let $O \in \mathcal{E}$. Then, by the density of $\mathbb{Q}$ in $\mathbb{R}$, for each $x \in O$ there exist $q_x, r_x \in \mathbb{Q}$ such that $x \in B(q_x, r_x) \subseteq O$ and hence $O = \bigcup_{x\in O} B(q_x, r_x)$. Since there are countably many pairs of the form (q_x, r_x), one can replace this union by a countable union, say, $O = \bigcup_{i\in\mathbb{N}} B(q_i, r_i)$

for some rationals $\{q_i, r_i\}_{i\in\mathbb{N}}$. Clearly $B(q_i, r_i) \in \mathcal{E}_2$ for every $i \in \mathbb{N}$ and hence $O \in \mathcal{M}(\mathcal{E}_2)$. Thus $\mathcal{E} \subseteq \mathcal{M}(\mathcal{E}_2)$ which implies that $\mathcal{B}(\mathbb{R}) = \mathcal{M}(\mathcal{E}) \subseteq \mathcal{M}(\mathcal{E}_2)$. Therefore $\mathcal{M}(\mathcal{E}_2) = \mathcal{B}(\mathbb{R})$.

We now show that $\mathcal{M}(\mathcal{E}_2) = \mathcal{M}(\mathcal{E}_1)$. Clearly $\mathcal{E}_2 \subseteq \mathcal{E}_1 \subseteq \mathcal{M}(\mathcal{E}_1)$ and hence $\mathcal{M}(\mathcal{E}_2) \subseteq \mathcal{M}(\mathcal{E}_1)$. To show the converse inclusion, let $(a, b) \in \mathcal{E}_1$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, we can find a decreasing sequence of rationals $(a_n)_{n\in\mathbb{N}}$ and an increasing sequence of rationals $(b_n)_{n\in\mathbb{N}}$ such that $a \le a_n, b_n \le b$ for all $n \in \mathbb{N}$, $\lim_{n\to\infty} a_n = a$ and $\lim_{n\to\infty} b_n = b$. Then $(a, b) = \bigcup_{n=1}^{\infty}(a_n, b_n)$. As $(a_n, b_n) \in \mathcal{E}_2$ for all $n \in \mathbb{N}$, we have that $(a, b) \in \mathcal{M}(\mathcal{E}_2)$. Thus $\mathcal{E}_1 \subseteq \mathcal{M}(\mathcal{E}_2)$ and hence $\mathcal{M}(\mathcal{E}_1) \subseteq \mathcal{M}(\mathcal{E}_2)$.

We will now prove that $\mathcal{M}(\mathcal{E}_3) = \mathcal{M}(\mathcal{E}_1)$. Let $[a, b] \in \mathcal{E}_3$. Then

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

Since the sets on the right-hand side are in $\mathcal{E}_1$, we have that $[a, b] \in \mathcal{M}(\mathcal{E}_1)$. Thus $\mathcal{E}_3 \subseteq \mathcal{M}(\mathcal{E}_1)$ and hence $\mathcal{M}(\mathcal{E}_3) \subseteq \mathcal{M}(\mathcal{E}_1)$. Let $(a, b) \in \mathcal{E}_1$. Then

$$(a, b) = \left(\bigcup_{n=1}^{\infty} [b, b + n] \cup [a - n, a] \right)^c$$

Since the sets on the right-hand side are in $\mathcal{E}_3$, we have that $(a, b) \in \mathcal{M}(\mathcal{E}_3)$. Thus $\mathcal{E}_1 \subseteq \mathcal{M}(\mathcal{E}_3)$ and hence $\mathcal{M}(\mathcal{E}_1) \subseteq \mathcal{M}(\mathcal{E}_3)$.

The rest of the proof is left to the reader as an exercise. □

Most "naturally occuring" sets in mathematics turn out to be Borel, even though this may not be immediately seen. We shall next give some examples of Borel subsets of Polish spaces that one runs into in practice.

- Let $f : \mathbb{R} \to \mathbb{R}$ be any function. The set

$$D_f = \{x \in \mathbb{R} : f \text{ is discontinuous at } x\}$$

  is a Borel set. This is because one can write this set as

$$\begin{aligned} D_f &= \{x \in \mathbb{R} : \exists \epsilon \in \mathbb{Q}^+ \ \forall \delta \in \mathbb{Q}^+ \ \exists y \in B(x, \delta) \ |f(y) - f(x)| \ge \epsilon\} \\ &= \{x \in \mathbb{R} : \exists \epsilon \in \mathbb{Q}^+ \ \forall \delta \in \mathbb{Q}^+ \ \exists y, z \in B(x, \delta) \ |f(y) - f(z)| \ge \epsilon\} \\ &= \bigcup_{\epsilon\in\mathbb{Q}^+} \bigcap_{\delta\in\mathbb{Q}^+} \{x \in \mathbb{R} : \exists y, z \in B(x, \delta) \ |f(y) - f(z)| \ge \epsilon\} \end{aligned}$$

  The reader is expected to check that, for each fixed $\epsilon, \delta \in \mathbb{Q}^+$, the innermost set is open. Therefore, D_f is a $\mathbf{\Sigma}^0_3$-subset of $\mathbb{R}$.
- The set C of convergent real sequences is a Borel subset of $\mathbb{R}^{\mathbb{N}}$ because one can write this set as

$$\begin{aligned} C &= \{(x_n) \in \mathbb{R}^{\mathbb{N}} : \forall \epsilon \in \mathbb{Q}^+ \exists k \in \mathbb{N} \ \forall i, j \ge k \ |x_i - x_j| < \epsilon\} \\ &= \bigcap_{\epsilon\in\mathbb{Q}^+} \bigcup_{k\in\mathbb{N}} \bigcap_{i\ge k} \bigcap_{j\ge k} \{(x_n) \in \mathbb{R}^{\mathbb{N}} : |x_i - x_j| < \epsilon\} \end{aligned}$$

For every fixed $\epsilon \in \mathbb{Q}^+$ and $i, j \in \mathbb{N}^+$, the inner-most set is closed in the product topology of $\mathbb{R}^{\mathbb{N}}$ and hence, C is a $\mathbf{\Pi}^0_3$-subset of $\mathbb{R}^{\mathbb{N}}$.

For more interesting examples of naturally occuring Borel sets, we refer the reader to [Kec95, Section 23].

1.4. **Product σ-algebras.** We shall call a pair $(\mathbf{X}, \mathcal{M})$ a *measurable space* if $\mathbf{X}$ is a non-empty set and $\mathcal{M}$ is a σ-algebra on $\mathbf{X}$. The sets in $\mathcal{M}$ are called the *measurable sets* of $(\mathbf{X}, \mathcal{M})$.

Given a collection $\{(\mathbf{X}_i, \mathcal{M}_i) : i \in I\}$ of indexed system of measurable spaces, we define their product (measurable) space to be the measurable space

$$\left(\prod_{i \in I} \mathbf{X}_i, \bigotimes_{i \in I} \mathcal{M}_i \right)$$

where $\bigotimes_{i \in I} \mathcal{M}_i$ is the *product σ-algebra* on $\prod_{i \in I} \mathbf{X}_i$ which is generated by the collection

$$\mathcal{E} = \{\pi_j^{-1}[A_j] : A_j \in \mathcal{M}_j,\ j \in I\}$$

and $\pi_j : \prod_{i \in I} \mathbf{X}_i \to \mathbf{X}_j$ are the projection maps for each $j \in I$. As an exercise, the reader is expected to check that if $\mathcal{M}_i$ are generated by $\mathcal{E}_i$, then the collection $\{\pi_j^{-1}[A_j] : A_j \in \mathcal{E}_j,\ j \in I\}$ also generates $\bigotimes_{i \in I} \mathcal{M}_i$.

In the case that the index set is countable, one can find another canonical generating set for the product σ-algebra. Suppose that I is countable and let

$$\hat{\mathcal{E}} = \left\{ \prod_{i \in I} A_i : A_i \in \mathcal{M}_i \right\}$$

Then $\mathcal{E} \subseteq \hat{\mathcal{E}} \subseteq \mathcal{M}(\hat{\mathcal{E}})$ and hence $\mathcal{M}(\mathcal{E}) \subseteq \mathcal{M}(\hat{\mathcal{E}})$. On the other hand, for any $\prod_{i \in I} A_i \in \hat{\mathcal{E}}$, we have that

$$\prod_{i \in I} A_i = \bigcap_{i \in I} \pi_i^{-1}[A_i] \in \mathcal{M}(\mathcal{E})$$

and hence $\mathcal{M}(\hat{\mathcal{E}}) \subseteq \mathcal{M}(\mathcal{E})$. Therefore, $\hat{\mathcal{E}}$ generates the product σ-algebra as well.

We conclude this subsection by proving that the Borel σ-algebra of the finite product of separable metric spaces has the same measurable sets as the product σ-algebra of the corresponding spaces endowed with their Borel σ-algebra.

Theorem 4. *Let $\mathbf{X}_i$ be a separable metric space for $1 \leq i \leq n$. Then we have $\bigotimes_{i=1}^n \mathcal{B}(\mathbf{X}_i) = \mathcal{B}\left(\prod_{i=1}^n \mathbf{X}_i\right)$ where $\prod_{i=1}^n \mathbf{X}_i$ is endowed with the product topology.*

Proof. As $\mathcal{B}(\mathbf{X}_i)$ are generated by the open sets of $\mathbf{X}_i$, we have that $\bigotimes_{i=1}^n \mathcal{B}(\mathbf{X}_i)$ is generated by $\{\pi_i^{-1}[U_i] : U_i \subseteq \mathbf{X}_i \text{ is open},\ 1 \leq i \leq n\}$. Recall that, by the definition of the product topology, these sets are open in $\prod_{i=1}^n \mathbf{X}_i$ and hence are in $\mathcal{B}\left(\prod_{i=1}^n \mathbf{X}_i\right)$. This shows that $\bigotimes_{i=1}^n \mathcal{B}(\mathbf{X}_i) \subseteq \mathcal{B}\left(\prod_{i=1}^n \mathbf{X}_i\right)$.

To show the converse inclusion, by the separability of $\mathbf{X}_i$, choose a countable dense set $D_i \subseteq \mathbf{X}_i$ for every $1 \leq i \leq n$. Consider the collections

$$\mathcal{E}_i = \{B(x,q) \subseteq \mathbf{X}_i : x \in D_i,\ q \in \mathbb{Q}^+\}$$

for every $1 \leq i \leq n$. It is easily seen that every open set of $\mathbf{X}_i$ is a countable union of the (open) sets in $\mathcal{E}_i$. On the other hand, the collection

$$\mathcal{U} = \left\{\prod_{i=1}^{n} U_i : U_i \subseteq \mathbf{X}_i \text{ is open}\right\}$$

is a basis for the product topology on $\prod_{i=1}^{n} \mathbf{X}_i$. Indeed, since a countable product of separable metric spaces is separable, every open set of $\prod_{i=1}^{n} \mathbf{X}_i$ can be written as a countable union of the elements of $\mathcal{U}$. It follows that every open set of $\prod_{i=1}^{n} \mathbf{X}_i$ is a countable union of the elements of

$$\left\{\prod_{i=1}^{n} A_i : A_i \in \mathcal{E}_i\right\}$$

These sets are clearly in $\bigotimes_{i=1}^{n} \mathcal{B}(\mathbf{X}_i)$ and hence the open sets of $\prod_{i=1}^{n} \mathbf{X}_i$ are in $\bigotimes_{i=1}^{n} \mathcal{B}(\mathbf{X}_i)$, which shows that $\mathcal{B}\left(\prod_{i=1}^{n} \mathbf{X}_i\right) \subseteq \bigotimes_{i=1}^{n} \mathcal{B}(\mathbf{X}_i)$. □

An immediate corollary of Theorem 4 is that $\mathcal{B}(\mathbb{R}^n) = \bigotimes_{i=1}^{n} \mathcal{B}(\mathbb{R})$ and hence the Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$ is generated by

$$\left\{\prod_{i=1}^{n} B_i : B_i \in \mathcal{B}(\mathbb{R})\right\}$$

and in fact, with some effort, it can be shown that $\mathcal{B}(\mathbb{R}^n)$ is generated by the collection

$$\left\{\prod_{i=1}^{n} (a_i, b_i) : a_i, b_i \in \mathbb{R}\right\}$$

1.5. **Exercises.** Below you shall find some exercises that you can work on regarding the topics in this section. These exercises are *not* to be handed in as homework assignments.

- Exercises 1, 4, 5 from Chapter 1 of [Fol99].
- Exercises 1, 5, 6, 9.a from Chapter 1.1 of [Coh93].
- Show that the following subset of $\mathbb{R}^2$ is Borel (and indeed, is F_σ.)

$$\left\{(x,y) \in \mathbb{R}^2 : xy \in \mathbb{Q}\right\}$$

2. Measures

2.1. **Definition, examples and properties.** Let $(\mathbf{X}, \mathcal{M})$ be a measurable space. A *(σ-additive) measure* on the measurable space $(\mathbf{X}, \mathcal{M})$ is a map $\mu : \mathcal{M} \to [0, +\infty]$ such that

- $\mu(\emptyset) = 0$ and
- If $A_1, A_2, \dots \in \mathcal{M}$ are disjoint, then $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$.

If the function μ satisfies finite additivity instead of σ-additivity, then one says that $\mu : \mathcal{M} \to [0, +\infty]$ is a *finitely additive measure* on $(\mathbf{X}, \mathcal{M})$.[1] From now on, unless specified otherwise, the word "measure" should be understood as σ-additive measure.

Given a (σ-additive or finitely additive) measure μ on $(\mathbf{X}, \mathcal{M})$, the triple $(\mathbf{X}, \mathcal{M}, \mu)$ is called a *measure space*. Next will be introduced some terminology for special types of measure spaces that are used frequently. A measure space $(\mathbf{X}, \mathcal{M}, \mu)$ is said to be a

- *probability space* if $\mu(\mathbf{X}) = 1$.
- *finite measure space* if $\mu(\mathbf{X}) < \infty$.
- *σ-finite measure space* if $\mathbf{X} = \bigcup_{i=1}^{\infty} X_i$ for some $X_i \in \mathcal{M}$ with $\mu(X_i) < \infty$.

Here are some examples of measure spaces.

- Let $\mathbf{X}$ be a non-empty countable set. The triple $(\mathbf{X}, \mathcal{P}(\mathbf{X}), \nu)$ is a measure space where $\nu : \mathcal{P}(\mathbf{X}) \to [0, \infty]$ is *the counting measure* given by
$$\nu(A) = \begin{cases} |A| & \text{if } A \text{ is finite} \\ +\infty & \text{otherwise} \end{cases}$$
- Let $\mathbf{X}$ be an uncountable set and $\mathcal{M} = \{A \subseteq \mathbf{X} : A \text{ or } A^c \text{ is countable}\}$. The triple $(\mathbf{X}, \mathcal{M}, \eta)$ is a probability space where $\eta : \mathcal{M} \to [0, \infty]$ is given by
$$\eta(A) = \begin{cases} 1 & \text{if } A^c \text{ is countable} \\ 0 & \text{if } A \text{ is countable} \end{cases}$$
- Let $\mathbf{X}$ be a non-empty set and $a \in \mathbf{X}$. The triple $(\mathbf{X}, \mathcal{P}(\mathbf{X}), \mu_a)$ is a probability space where $\mu_a : \mathcal{P}(\mathbf{X}) \to [0, \infty]$ is the *Dirac measure at a* given by
$$\mu_a(S) = \begin{cases} 1 & \text{if } a \in S \\ 0 & \text{if } a \notin S \end{cases}$$
- Let $\mathbf{X}$ be a non-empty countable set and $f : \mathbf{X} \to [0, \infty]$ be any function. The triple $(\mathbf{X}, \mathcal{P}(\mathbf{X}), \mu)$ is a measure space where $\mu : \mathcal{P}(\mathbf{X}) \to [0, \infty]$ is given by
$$\mu(A) = \sum_{x \in A} f(x)$$
- Let $\mathcal{M} = \{\emptyset, 2\mathbb{N}, 2\mathbb{N} + 1, \mathbb{N}\}$. The triple $(\mathbb{N}, \mathcal{M}, \xi)$ is a probability space where $\xi(\emptyset) = \xi(2\mathbb{N} + 1) = 0$ and $\xi(2\mathbb{N}) = \xi(\mathbb{N}) = 1$.

We shall now prove some basic properties of measures that easily follow from the definition.

---

[1]We would like to note that, while working with finitely additive measures, one may only require $\mathcal{M}$ to be an algebra instead of a σ-algebra for it suffices for a domain of a finitely additive measure to be closed under finite unions.

Theorem 5. *Let $(\mathbf{X}, \mathcal{M}, \mu)$ be a measure space. Then*

a. *If $A, B \in \mathcal{M}$ and $A \subseteq B$, then $\mu(A) \leq \mu(B)$.*
b. *If $A_1, A_2, \cdots \in \mathcal{M}$, then $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu(A_i)$.*
c. *If $A_1, A_2, \cdots \in \mathcal{M}$ and $A_i \subseteq A_{i+1}$ for every $i \in \mathbb{N}^+$, then*

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} \mu(A_i)$$

d. *If $A_1, A_2, \cdots \in \mathcal{M}$ and $A_i \supseteq A_{i+1}$ for every $i \in \mathbb{N}^+$ and $\mu(A_1) < \infty$, then*

$$\mu\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} \mu(A_i)$$

Proof. To prove (a), let $A, B \in \mathcal{M}$ be with $A \subseteq B$. Then $B = A \sqcup (B - A)$ and hence $\mu(A) \leq \mu(A) + \mu(B - A) = \mu(A \sqcup (B - A)) = \mu(B)$.

To prove (b), let $A_1, A_2, \cdots \in \mathcal{M}$. Set $B_1 = A_1$ and $B_{i+1} = A_{i+1} - \bigcup_{k=1}^{i} A_k$ for all $i \in \mathbb{N}^+$. Then $B_i \in \mathcal{M}$ for all $i \in \mathbb{N}^+$ and moreover, B_i's are disjoint. It follows that

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \mu\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} \mu(B_i) \leq \sum_{i=1}^{\infty} \mu(A_i)$$

To prove (c), let $A_1, A_2, \cdots \in \mathcal{M}$ be such that $A_i \subseteq A_{i+1}$ for every $i \in \mathbb{N}^+$. Then

$$\begin{aligned} \mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \mu\left(A_1 \sqcup \bigsqcup_{i=1}^{\infty} (A_{i+1} - A_i)\right) &= \mu(A_1) + \sum_{i=1}^{\infty} \mu(A_{i+1} - A_i) = \\ &= \lim_{n\to\infty} \mu(A_1) + \sum_{i=1}^{n-1} \mu(A_{i+1} - A_i) \\ &= \lim_{n\to\infty} \mu\left(A_1 \sqcup \bigsqcup_{i=1}^{n} (A_{i+1} - A_i)\right) \\ &= \lim_{n\to\infty} \mu(A_n) \end{aligned}$$

To prove (d), let $A_1, A_2, \cdots \in \mathcal{M}$ be such that $\mu(A_1) < \infty$ and $A_i \supseteq A_{i+1}$ for every $i \in \mathbb{N}^+$. Set $B_i = A_1 - A_i$. Then $B_i \in \mathcal{M}$ and $B_i \subseteq B_{i+1}$ and $A_1 = A_i \sqcup B_i$ for every $i \in \mathbb{N}^+$. Moreover, $\bigcup_{i=1}^{\infty} B_i \sqcup \bigcap_{i=1}^{\infty} A_i = A_1$. It follows from part (c) that

$$\begin{aligned} \mu(A_1) = \mu\left(\bigcup_{i=1}^{\infty} B_i \sqcup \bigcap_{i=1}^{\infty} A_i\right) &= \mu\left(\bigcup_{i=1}^{\infty} B_i\right) + \mu\left(\bigcap_{i=1}^{\infty} A_i\right) \\ &= \left(\lim_{n\to\infty} \mu(B_n)\right) + \mu\left(\bigcap_{i=1}^{\infty} A_i\right) \\ &= \left(\lim_{n\to\infty} \mu(A_1) - \mu(A_n)\right) + \mu\left(\bigcap_{i=1}^{\infty} A_i\right) \end{aligned}$$

Substracting $\mu(A_1)$ from both sides, we get that $\lim_{n\to\infty} \mu(A_n) = \mu\left(\bigcap_{i=1}^{\infty} A_i\right)$. ☐

Let $(\mathbf{X}, \mathcal{M}, \mu)$ be a measure space. A measurable set $A \in \mathcal{M}$ is said to be μ-*null* if $\mu(A) = 0$. A statement quantifying over the points of the measure space $(\mathbf{X}, \mathcal{M}, \mu)$ is said to hold μ-*almost everywhere* if the set of points where it fails is a μ-null set. If the measure μ is understood from the context, we shall simply write *null* and *almost everywhere* (or, a.e.)

For example, consider the measure space $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$ where μ is the Dirac measure concentrated at 7. Let $f : \mathbb{N} \to \mathbb{N}$ be the identity function. Then the statement "$f(n) = 7$ almost everywhere." is true since

$$\mu_7 (\{n \in \mathbb{N} : f(n) \neq 7\}) = \mu_7 (\mathbb{N} - \{7\}) = 0$$

2.2. **Null sets and completing measures.** Let $(\mathbf{X}, \mathcal{M}, \mu)$ be a measure space and set

$$\mathcal{M}_{null} = \{A \in \mathcal{M} : A \textit{ is null }\} = \{A \in \mathcal{M} : \mu(A) = 0\}$$

It is easily seen that $\mathcal{M}_{null}$ is closed under countable unions and that the monotonicity of measures implies that $B \in \mathcal{M}_{null}$ whenever $A \in \mathcal{M}_{null}$ and $B \subseteq A$ and $B \in \mathcal{M}$. On the other hand, $\mathcal{M}_{null}$ need not be closed under taking arbitrary subsets unless the taken subset is already in $\mathcal{M}$. For technical reasons, we often want to have all subsets of null sets to be measurable and consequently, null themselves.

The measure space $(\mathbf{X}, \mathcal{M}, \mu)$ is said to be *complete* if $A \in \mathcal{M}_{null}$ and $B \subseteq A$ implies that $B \in \mathcal{M}$ (and consequently, $B \in \mathcal{M}_{null}$.) For example, any measure space equipped with the counting measure is complete, whereas, the measure space $(\mathbb{N}, \{\emptyset, 2\mathbb{N}, 2\mathbb{N}+1, \mathbb{N}\}, \xi)$ where $\xi(\emptyset) = \xi(2\mathbb{N}+1) = 0$ and $\xi(2\mathbb{N}) = \xi(\mathbb{N}) = 1$ is not complete.

It turns out that any measure space can be completed simply by adding the subsets of its null sets to its σ-algebra and extending the measure appropriately.

Theorem 6. *Let $(\mathbf{X}, \mathcal{M}, \mu)$ be a measure space. Then there exists a complete measure space $(\mathbf{X}, \overline{\mathcal{M}}, \overline{\mu})$ such that $\mathcal{M} \subseteq \overline{\mathcal{M}}$ and $\overline{\mu} \upharpoonright \mathcal{M} = \mu$.*

Proof. Let $\overline{\mathcal{M}} = \{A \cup B : A \in \mathcal{M}, B \subseteq N \text{ for some } N \in \mathcal{M}_{null}\}$. Clearly $\mathcal{M} \subseteq \overline{\mathcal{M}}$. We claim that $\overline{\mathcal{M}}$ is a σ-algebra on $\mathbf{X}$.

Let $C_1, C_2, \dots \in \overline{\mathcal{M}}$. Then, by definition, for every $i \in \mathbb{N}^+$, we have that $C_i = A_i \cup B_i$ for some $A_i \in \mathcal{M}$ and $N_i \in \mathcal{M}_{null}$ with $B_i \subseteq N_i$. Note that, $\mathcal{M}$ and $\mathcal{M}_{null}$ being closed under countable unions implies that $\bigcup_{i=1}^{\infty} A_i \in \mathcal{M}$ and $\bigcup_{i=1}^{\infty} B_i \subseteq \bigcup_{i=1}^{\infty} N_i \in \mathcal{M}_{null}$. It follows that

$$\bigcup_{i=1}^{\infty} C_i = \bigcup_{i=1}^{\infty} (A_i \cup B_i) = \bigcup_{i=1}^{\infty} A_i \cup \bigcup_{i=1}^{\infty} B_i \in \overline{\mathcal{M}}$$

Thus $\overline{M}$ is closed under countable unions. Let $C \in \overline{\mathcal{M}}$. Then $C = A \cup B$ for some $A \in \mathcal{M}$ and $N \in \mathcal{M}_{null}$ with $B \subseteq N$. Then

$$C = A \cup B = A \cup (B - A) = (A \cup (N - A)) \cap ((N - A)^c \cup B)$$

and hence

$$C^c = (A \cup B)^c = (A \cup (B - A))^c = (A \cup (N - A))^c \cup ((N - A) - B)$$

Note that, as $\mathcal{M}$ is a σ-algebra and $A, N \in \mathcal{M}$, we have $(A \cup (N - A))^c \in \mathcal{M}$ and moreover, $((N - A) - B) \subseteq N \in \mathcal{M}_{null}$. It follows that $C^c \in \overline{\mathcal{M}}$. Thus $\overline{\mathcal{M}}$ is closed under complementation, completing the proof that $\overline{\mathcal{M}}$ is a σ-algebra.

We shall now define a measure $\overline{\mu}$ on $\overline{\mathcal{M}}$. Consider the relation $\overline{\mu}$ given by

$$\overline{\mu}(C) = \mu(A) \Leftrightarrow \exists A \in \mathcal{M}\ \exists N \in \mathcal{M}_{null}\ C = A \cup B\ \wedge\ B \subseteq N$$

We claim that $\overline{\mu}$ is well-defined and hence is indeed a map from $\overline{\mu}$ to $[0, \infty]$. Assume that $C = A \cup B = A' \cup B'$ for some $A, A' \in \mathcal{M}$ and $N, N' \in \mathcal{M}_{null}$ with $B \subseteq N$ and $B' \subseteq N'$. Then, since $A \subseteq A' \cup B' \subseteq A' \cup N'$ and $A' \subseteq A \cup B \subseteq A \cup N$, by the monotonicity of μ, we have that

$$\mu(A) \leq \mu(A') + \mu(N') = \mu(A') \leq \mu(A) + \mu(N) = \mu(A)$$

Therefore, $\overline{\mu}(C)$ is equal to $\mu(A)$ for any $A \in \mathcal{M}$ which is equal to C modulo a subset of a null set. This shows that $\overline{\mu}$ is well-defined.

We now check that $\overline{\mu}$ is indeed a measure. That $\overline{\mu}(\emptyset) = 0$ is trivial. To show σ-additivity, let $C_1, C_2, \cdots \in \overline{\mathcal{M}}$ be disjoint. Then, for each $i \in \mathbb{N}^+$, there exist $A_i \in \mathcal{M}$ and $B_i \subseteq N_i \in \mathcal{M}_{null}$ such that $C_i = A_i \cup B_i$. It is clear that A_i's are disjoint and that $\bigcup_{i=1}^{\infty} C_i = \bigcup_{i=1}^{\infty} A_i \ \cup \ \bigcup_{i=1}^{\infty} B_i$. On the other hand, since $\bigcup_{i=1}^{\infty} A_i \in \mathcal{M}$ and $\bigcup_{i=1}^{\infty} B_i \subseteq \bigcup_{i=1}^{\infty} N_i \in \mathcal{M}_{null}$, by definition, we have

$$\overline{\mu}\left(\bigcup_{i=1}^{\infty} C_i\right) = \mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i) = \sum_{i=1}^{\infty} \overline{\mu}(C_i)$$

That $\overline{\mu} \restriction \mathcal{M} = \mu$ and that $(\mathbf{X}, \overline{\mathcal{M}}, \overline{\mu})$ is complete are left to the reader to be checked. □

Recall the definition of the complete measure space $(\mathbf{X}, \overline{\mathcal{M}}, \overline{\mu})$ constructed in the previous proof. It is an exercise to the reader to check that

- If $(\mathbf{X}, \overline{\mathcal{M}}, \mu')$ is a complete measure space with $\mu' \restriction \mathcal{M} = \mu$, then $\mu' = \overline{\mu}$.
- If $(\mathbf{X}, \mathcal{M}', \mu')$ is a complete measure space with $\mathcal{M} \subseteq \mathcal{M}'$ and $\mu' \restriction \mathcal{M} = \mu$, then $\overline{\mathcal{M}} \subseteq \mathcal{M}'$.

In other words, $(\mathbf{X}, \overline{\mathcal{M}}, \overline{\mu})$ is the "smallest" complete measure space extending $(\mathbf{X}, \mathcal{M}, \mu)$. The measure space $(\mathbf{X}, \overline{\mathcal{M}}, \overline{\mu})$ is called the *completion* of $(\mathbf{X}, \mathcal{M}, \mu)$.

We would like to note that not every complete measure space extending a measure space is the completion. For example, the measure space $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu_2)$ where μ_2 is the Dirac measure concentrated at 2 is a complete measure space which extends the (incomplete) measure space $(\mathbb{N}, \{\emptyset, 2\mathbb{N}, 2\mathbb{N}+1, \mathbb{N}\}, \xi)$ where $\xi(\emptyset) = \xi(2\mathbb{N}+1) = 0$ and $\xi(2\mathbb{N}) = \xi(\mathbb{N}) = 1$. On the other hand, as constructed in the proof of Theorem

6, the completion of this measure space is

$$(\mathbb{N}, \{A \cup B : A \in \{\emptyset, 2\mathbb{N}, 2\mathbb{N}+1, \mathbb{N}\}, B \subseteq 2\mathbb{N}+1\}, \overline{\xi})$$

where

$$\overline{\xi}(S) = \begin{cases} 1 & \text{if } 2\mathbb{N} \subseteq S \\ 0 & \text{otherwise} \end{cases}$$

2.3. **Outer measures.** In this subsection, we shall introduce the notion of an outer measure, which will be later used to construct measures on σ-algebras that are extending pre-specified functions defined on algebras.

Let $\mathbf{X}$ be a non-empty set. A function $\mu^* : \mathcal{P}(\mathbf{X}) \to [0, \infty]$ is said to be an *outer measure* on $\mathbf{X}$ if

- $\mu^*(\emptyset) = 0$,
- If $A \subseteq B \subseteq \mathbf{X}$, then $\mu^*(A) \leq \mu^*(B)$, and,
- If $A_1, A_2, \cdots \subseteq \mathbf{X}$, then $\mu^*\left(\bigcup_{i=1}^\infty A_i\right) \leq \sum_{i=1}^\infty \mu^*(A_i)$.

One way to construct outer measures is to pre-define outer measures of certain "elementary sets" that can cover the whole set and define the outer measure of a subset as the infimum of the sums of outer measures of elementary sets that cover it. More specifically, we have the following.

Theorem 7. *Let $\mathcal{E} \subseteq \mathcal{P}(\mathbf{X})$ be such that $\emptyset, \mathbf{X} \in \mathcal{E}$ and let $\rho : \mathcal{E} \to [0, \infty]$ be an arbitrary function with $\rho(\emptyset) = 0$. Then the map $\mu^* : \mathcal{P}(\mathbf{X}) \to [0, \infty]$ given by*

$$\mu^*(A) = \inf \left\{ \sum_{i=1}^\infty \rho(E_i) : E_i \in \mathcal{E},\ A \subseteq \bigcup_{i=1}^\infty E_i \right\}$$

is an outer measure.

Proof. That $\mu^*(\emptyset) = 0$ follows from that $\emptyset \in \mathcal{E}$ and $\rho(\emptyset) = 0$. That $\mu^*(A) \leq \mu^*(B)$ whenever $A \subseteq B$ follows from the fact that any covering of B is a covering of A whenever $A \subseteq B$. It remains to check that μ^* is countably subadditive, that is, it satisfies the third condition in the definition of an outer measure.

Let $A_1, A_2, \cdots \subseteq \mathbf{X}$ and $\epsilon > 0$. Then, by definition, for every $i \in \mathbb{N}^+$, we can find $E_1^i, E_2^i, \cdots \in \mathcal{E}$ such that $A_i \subseteq \bigcup_{k=1}^\infty E_k^i$ and

$$\sum_{k=1}^\infty \rho(E_k^i) \leq \mu^*(A_i) + \frac{\epsilon}{2^i}$$

Since we have $\bigcup_{i=1}^\infty A_i \subseteq \bigcup_{i,k=1}^\infty E_k^i$, we have that

$$\mu^*\left(\bigcup_{i=1}^\infty A_i\right) \leq \sum_{i,k=1}^\infty \rho(E_k^i) \leq \sum_{i=1}^\infty \left(\mu^*(A_i) + \frac{\epsilon}{2^i}\right) \leq \left(\sum_{i=1}^\infty \mu^*(A_i)\right) + \epsilon$$

Since this inequality is true for every $\epsilon > 0$, we have that

$$\mu^*\left(\bigcup_{i=1}^\infty A_i\right) \leq \sum_{i=1}^\infty \mu^*(A_i)$$

Thus μ^* is an outer measure on $\mathbf{X}$. $\square$

We shall now learn how to pass from outer measures to (complete) measures. Let μ^* be an outer measure on $\mathbf{X}$. A subset $A \subseteq \mathbf{X}$ is said to be μ^**-measurable* if

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

for all $E \subseteq \mathbf{X}$. We will prove soon that the set of μ^*-measurable subsets of $\mathbf{X}$ is a σ-algebra which, together with μ^*, will form a complete measure space on $\mathbf{X}$. What is the intuition behind this magically-working technical condition, which is due to Carathéodory?

One can think of the value $\mu^*(E) - \mu^*(E \cap A^c)$ as the "inner measure" of the set $E \cap A$. With this interpretation, Carathéodory's conditions simply says that A is μ^*-measurable if and only if the outer measure of $E \cap A$ is equal to the "inner measure" of $E \cap A$ for all $E \subseteq \mathbf{X}$. In a sense, a set A is μ^*-measurable if the outer measure is equal to the "inner measure" for every possible "slice" of A.

Let us note several observations regarding Carathéodory's condition. First, the inequality

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

holds since μ^* is subadditive. Thus, in order to show that Carathéodory's condition holds for a set A, it suffices to prove the converse inequality

$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

for all $E \subseteq \mathbf{X}$. Second, this latter inequality trivially holds for all $E \subseteq \mathbf{X}$ with $\mu^*(E) = \infty$. Thus, Carathéodory's condition holds for a set A if and only if

$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

for all $E \subseteq \mathbf{X}$ with $\mu^*(E) < \infty$. We are now ready to state and prove our main theorem.

Theorem 8. *Let $\mathbf{X}$ be a non-empty set, $\mu^* : \mathcal{P}(\mathbf{X}) \to [0, \infty]$ be an outer measure and let $\mathcal{M} = \{A \subseteq \mathbf{X} : A \text{ is } \mu^*\text{-measurable}\}$. Then $(\mathbf{X}, \mathcal{M}, \mu^* \restriction \mathcal{M})$ is a complete measure space.*

Proof. We will first prove that $\mathcal{M}$ is a σ-algebra. As a first step, we shall show that $\mathcal{M}$ is an algebra. That $\mathcal{M}$ is closed under complementation is trivial since Carathéodory's condition is symmetric with respect to A and A^c. It remains to prove that $\mathcal{M}$ is closed under finite unions. Let $A, B \in \mathcal{M}$. Then, since A and B satisfies Carathéodory's condition, we have that

$$\begin{aligned}\mu^*(E) &= \mu^*(E \cap A) + \mu^*(E - A) \\ &= (\mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c)) + (\mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c))\end{aligned}$$

for all $E \subseteq \mathbf{X}$. On the other hand, $A \cup B = (A^c \cap B) \cup (A \cap B) \cup (A \cap B^c)$ and hence, by the subadditivity of μ^*, we have that

$$\mu^*(E \cap (A \cup B)) \leq \mu^*(E \cap (A^c \cap B)) + \mu^*(E \cap (A \cap B)) + \mu^*(E \cap (A \cap B^c))$$

for all $E \subseteq \mathbf{X}$. Combining these, we get that

$$\mu^*(E \cap (A \cup B)) \leq \mu^*(E) - \mu^*(E \cap A^c \cap B^c)$$

and hence

$$\mu^*(E) \geq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c) \geq \mu^*(E)$$

for all $E \subseteq \mathbf{X}$. Thus $A \cup B \in \mathcal{M}$ and hence, $\mathcal{M}$ is an algebra.

Having observed that $\mathcal{M}$ is an algebra, in order to show that $\mathcal{M}$ is closed under countable unions, it suffices to show that $\mathcal{M}$ is closed under countable disjoint unions. To that end, let $A_1, A_2, \cdots \in \mathcal{M}$ be disjoint. Set $B_n = A_1 \cup \cdots \cup A_n$ for every $n \geq 1$. We are going to prove by induction that

$$\mu^*(E \cap B_n) = \sum_{k=1}^{n} \mu^*(E \cap A_k)$$

for every $E \subseteq \mathbf{X}$ and for every $n \geq 1$. The claim trivially holds for $n = 1$. Let $n \geq 1$ and assume that the claim holds for n. As A_{n+1} is μ^*-measurable, we have that

$$\begin{aligned}
\mu^*(E \cap B_{n+1}) &= \mu^*((E \cap B_{n+1}) \cap A_{n+1}) + \mu^*((E \cap B_{n+1}) \cap A_{n+1}^c) \\
&= \mu^*(E \cap A_{n+1}) + \mu^*(E \cap B_n) \\
&= \mu^*(E \cap A_{n+1}) + \sum_{k=1}^{n} \mu^*(E \cap A_k) = \sum_{k=1}^{n+1} \mu^*(E \cap A_k)
\end{aligned}$$

for every $E \subseteq \mathbf{X}$. Therefore, by induction, the claim holds for all $n \geq 1$. Since $\mathcal{M}$ is an algebra, B_i's are also μ^*-measurable. It follows that

$$\mu^*(E) = \mu^*(E \cap B_n^c) + \mu^*(E \cap B_n) \geq \mu^*\left(E \cap \bigcap_{k=1}^{\infty} A_k^c\right) + \sum_{k=1}^{n} \mu^*(E \cap A_k)$$

for every $E \subseteq \mathbf{X}$ and for every $n \geq 1$. By taking limit, we get that

$$\begin{aligned}
\mu^*(E) &\geq \mu^*\left(E \cap \bigcap_{k=1}^{\infty} A_k^c\right) + \sum_{k=1}^{\infty} \mu^*(E \cap A_k) \\
&\geq \mu^*\left(E \cap \left(\bigcup_{k=1}^{\infty} A_k\right)^c\right) + \mu^*\left(E \cap \left(\bigcup_{k=1}^{\infty} A_k\right)\right) \geq \mu^*(E)
\end{aligned}$$

for every $E \subseteq \mathbf{X}$. It follows that $\bigcup_{k=1}^{\infty} A_k \in \mathcal{M}$ and hence $\mathcal{M}$ is a σ-algebra. Moreover, if one takes $E = \bigcup_{k=1}^{\infty} A_k$ in the previous inequality, one gets that

$$\sum_{k=1}^{\infty} \mu^*(A_k) = \sum_{k=1}^{\infty} \mu^*(E \cap A_k) = \mu^*\left(\bigcup_{k=1}^{\infty} A_k\right)$$

It follows that μ^* is countably additive on $\mathcal{M}$ and hence, the triple $(\mathbf{X}, \mathcal{M}, \mu* \restriction \mathcal{M})$ is a measure space. To show its completeness, let $A \in \mathcal{M}$ be with $\mu^*(A) = 0$ and let $B \subseteq A$. Then, $\mu^*(B) = 0$ by the monotonicity of μ^* and consequently, we have that

$$\mu^*(E) \leq \mu^*(E \cap B) + \mu^*(E \cap B^c) \leq \mu^*(B) + \mu^*(E \cap B^c) \leq \mu^*(E)$$

for every $E \subseteq \mathbf{X}$ and hence, $B \in \mathcal{M}$. This completes the proof that $(\mathbf{X}, \mathcal{M}, \mu* \restriction \mathcal{M})$ is a complete measure space. $\square$

As an application of Theorem 8, we will next develop a fundamental tool to construct measures on σ-algebras that are extending pre-specified functions defined on algebras, known as Carathéodory's extension theorem. Before we state this fundamental theorem, we need to introduce the notion of a premeasure on a set.

Let $\mathbf{X}$ be non-empty and $\mathcal{A} \subseteq \mathcal{P}(\mathbf{X})$ be an algebra on $\mathbf{X}$. A map $\rho : \mathcal{A} \to [0, \infty]$ is called a *premeasure* on $\mathcal{A}$ if

- $\rho(\emptyset) = 0$ and
- If $A_1, A_2, \dots \in \mathcal{A}$ with $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$, then $\rho\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \rho(A_i)$.

A premeasure behaves just like a measure, except that it is defined on an algebra which may not be a σ-algebra and consequently, the countable additivity condition has to be modified appropriately.

Theorem 9 (Carathéodory's extension theorem)**.** *Let $\mathbf{X}$ be non-empty, $\mathcal{A}$ be an algebra on $\mathbf{X}$ and $\rho : \mathcal{A} \to [0, \infty]$ be a premeasure on $\mathcal{A}$. Then there exists a measure $\mu : \mathcal{M}(\mathcal{A}) \to [0, \infty]$ such that $\mu \restriction \mathcal{A} = \rho$ where $\mathcal{M}(\mathcal{A})$ denotes the σ-algebra generated by $\mathcal{A}$.*

Proof. By Theorem 7, since $\emptyset, \mathbf{X} \in \mathcal{A}$ and $\rho(\emptyset) = 0$, the map $\mu^* : \mathcal{P}(\mathbf{X}) \to [0, \infty]$ given by

$$\mu^*(S) = \inf \left\{ \sum_{i=1}^{\infty} \rho(A_i) : A_i \in \mathcal{A},\ S \subseteq \bigcup_{i=1}^{\infty} A_i \right\}$$

is an outer measure on $\mathbf{X}$. By Theorem 8, the set $\mathcal{M}$ consisting of μ^*-measurable subsets of $\mathbf{X}$ is a σ-algebra on $\mathbf{X}$ and moreover, μ^* is countably additive on $\mathcal{M}$.

We will now prove that $\mathcal{A} \subseteq \mathcal{M}$. Let $A \in \mathcal{A}$, $E \subseteq \mathbf{X}$ and $\epsilon > 0$. Then, by definition of μ^*, there exist $A_1, A_2, \dots \in \mathcal{A}$ such that $\sum_{i=1}^{\infty} \rho(A_i) \leq \mu^*(E) + \epsilon$ and

$E \subseteq \bigcup_{i=1}^\infty A_i$. It follows that

$$\begin{aligned}\mu^*(E) &\leq \mu^*(E \cap A) + \mu^*(E \cap A^c) \\ &\leq \mu^*\left(\bigcup_{i=1}^\infty (A_i \cap A)\right) + \mu^*\left(\bigcup_{i=1}^\infty (A_i \cap A^c)\right) \\ &\leq \sum_{i=1}^\infty \rho(A_i \cap A) + \sum_{i=1}^\infty \rho(A_i \cap A^c) \\ &\leq \sum_{i=1}^\infty \rho(A_i) \leq \mu^*(E) + \epsilon\end{aligned}$$

Since $\epsilon > 0$ was arbitrary, we have that

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

Hence A is μ^*-measurable and so $\mathcal{A} \subseteq \mathcal{M}$, which implies $\mathcal{M}(\mathcal{A}) \subseteq \mathcal{M}$ because $\mathcal{M}$ is a σ-algebra.

Set $\mu = \mu^* \restriction \mathcal{M}(\mathcal{A})$. We claim that μ is as claimed. That μ is σ-additive follows from that μ^* is σ-additive on $\mathcal{M}$.

It remains to show that $\mu \restriction \mathcal{A} = \rho$. To see this, let $A \in \mathcal{A}$. Let $A_1, A_2, \dots \in \mathcal{A}$ be sets with $A \subseteq \bigcup_{i=1}^\infty A_i$. Define $B_1 = A \cap A_1$ and $B_{i+1} = A \cap \left(A_{i+1} - \bigcup_{k=1}^i A_i\right)$ for every $i \geq 1$. Then B_i's are disjoint and are in $\mathcal{A}$ and $A = \bigcup_{i=1}^\infty B_i$. It follows from the definition of μ^* that

$$\rho^*(A) = \rho^*\left(\bigcup_{i=1}^\infty B_i\right) = \sum_{i=1}^\infty \rho(B_i) \leq \sum_{i=1}^\infty \rho(A_i)$$

Therefore we have $\rho(A) \leq \mu^*(A)$. Since $A = A \cup \emptyset \cup \emptyset \cup \dots$, we clearly have $\mu^*(A) \leq \rho(A) + \rho(\emptyset) + \rho(\emptyset) + \dots = \rho(A)$. Thus $\mu^*(A) = \rho(A)$, which completes the proof that $\mu^* \restriction \mathcal{A} = \rho$. □

We would like to note two important points regarding Carathéodory's extension theorem. First, given an algebra $\mathcal{A}$ and a premeasure $\rho : \mathcal{A} \to [0, \infty]$, we not only have the measure space

$$(\mathbf{X}, \mathcal{M}(\mathcal{A}), \mu^* \restriction \mathcal{M}(\mathcal{A}))$$

with $\mu^* \restriction \mathcal{A} = \rho$ but indeed have a *complete* measure space

$$(\mathbf{X}, \mathcal{M}, \mu^* \restriction \mathcal{M})$$

where $\mathcal{M}$ is the set of μ^*-measurable sets and μ^* is the outer measure derived from ρ as defined in the proof of Theorem 9. As it has been noted in the proof, the relationship between these two σ-algebras is that $\mathcal{M}(\mathcal{A}) \subseteq \mathcal{M}$. Indeed, the latter measure space is simply the completion of the former whenever ρ is σ-finite.[2] The reader is expected to solve the exercise [Fol99, Exercise 1.18] from which this last

[2]The definition of σ-finiteness for premeasure is the same as that of measure, that is, we say that a premeasure ρ on $\mathcal{A} \subseteq \mathcal{P}(\mathbf{X})$ is σ-finite if $\mathbf{X} = \bigcup_{i=1}^\infty A_i$ for some $A_i \in \mathcal{A}$ with $\rho(A_i) < \infty$.

claim follows. Second, if ρ is σ-finite, then the extension $\mu = \mu^* \restriction \mathcal{M}$ given by Theorem 9 is indeed unique. For a proof of this fact, which will be later used but not proven here, we refer the reader to [Fol99, Theorem 1.14].

2.4. **Borel measures on $\mathbb{R}$.** Having developed a flexible tool to build measures from premeasures, we shall now use this tool to construct various measures on the Borel σ-algebra of $\mathbb{R}$. From now on, any measure defined on the measurable space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ will be called a *Borel measure* on $\mathbb{R}$.

In order to motivate our to-be-carried-out construction, let μ be a Borel measure on $\mathbb{R}$ which takes finite values on bounded Borel subsets of $\mathbb{R}$. Consider the function

$$F(x) = \begin{cases} \mu((0,x]) & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -\mu((x,0]) & \text{if } x < 0 \end{cases}$$

Then it follows from the properties of a measure that F is increasing[3] and right-continuous. It turns out that this process can be reversed and that any increasing right-continuous function also induces a Borel measure on $\mathbb{R}$ which takes finite values on bounded Borel subsets of $\mathbb{R}$. More precisely, we have the following theorem.

Theorem 10. *Let $F : \mathbb{R} \to \mathbb{R}$ be an increasing right-continuous function. Then there exists a (unique) Borel measure μ_F on $\mathbb{R}$ such that $\mu_F((a,b]) = F(b) - F(a)$ for all $a < b$.*

While we are not planning to prove this fact, whose proof is not conceptually difficult but is lengthy, we shall briefly describe how the proof goes. Let $\mathcal{A}$ be the collection of finite disjoint unions of sets of the form $(a,b]$, (b,∞) and $(-\infty,a]$ with $-\infty < a \leq b < \infty$. One can check that $\mathcal{A}$ is indeed an algebra on $\mathbb{R}$. Consider the map $\rho : \mathcal{A} \to [0,\infty]$ given by

$$\rho\left(\bigsqcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \hat{\rho}(A_i)$$

where

$$\hat{\rho}\left(A_i\right) = \begin{cases} F(b) - F(a) & \text{if } A_i = (a,b] \\ \lim\limits_{x\to\infty} F(x) - F(b) & \text{if } A_i = (b,\infty) \\ F(a) - \lim\limits_{x\to-\infty} F(x) & \text{if } A_i = (-\infty,a] \end{cases}$$

The crux of the matter is proving that ρ is a premeasure on $\mathcal{A}$. (The reader may refer to [Fol99, Proposition 1.15] for a proof of this.) Once this is proven, the rest is taken care of by Carathéodory's extension theorem. Since $\mathcal{M}(\mathcal{A}) = \mathcal{B}(\mathbb{R})$,

[3] Since we are following Folland's terminology, when we say "increasing", what we really mean is "non-decreasing", that is, $F(x) \leq F(y)$ whenever $x \leq y$.

Theorem 9 implies that there exists a Borel measure μ_F on $\mathbb{R}$ such that $\mu_F \restriction \mathcal{A} = \rho$. Moreover, by construction of this measure, we have that

$$\begin{aligned}\mu_F(A) &= \inf\left\{\sum_{i=1}^{\infty}\rho(S_i) : S_i \in \mathcal{A},\ A \subseteq \bigcup_{i=1}^{\infty} S_i\right\} \\ &= \inf\left\{\sum_{i=1}^{\infty}\hat{\rho}(A_i) : A \subseteq \bigcup_{i=1}^{\infty} A_i,\ A_i \text{ is of the form } (-\infty, a], (a, b] \text{ or } (b, \infty)\right\}\end{aligned}$$

Notice that we have $(a, \infty) = \bigsqcup_{i=0}^{\infty}(a+i, a+i+1]$ and $(-\infty, a] = \bigsqcup_{i=0}^{\infty}(a-i-1, a-i]$. Moreover, we have that

$$\hat{\rho}((a,\infty)) = \lim_{x\to\infty} F(x) - F(a) = \sum_{i=0}^{\infty} F(a+i+1) - F(a+i) = \sum_{i=0}^{\infty}\hat{\rho}((a+i, a+i+1])$$

and

$$\hat{\rho}((-\infty, a]) = F(a) - \lim_{x\to-\infty} F(x) = \sum_{i=0}^{\infty} F(a-i) - F(a-i-1) = \sum_{i=0}^{\infty}\hat{\rho}((a-i-1, a-i])$$

It follows that, in the definition of μ_F, it suffices to consider the bounded half-open intervals in $\mathcal{A}$ and hence we have that

$$\begin{aligned}\mu_F(A) &= \inf\left\{\sum_{i=1}^{\infty}\hat{\rho}((a_i, b_i]) : A \subseteq \bigcup_{i=1}^{\infty}(a_i, b_i]\right\} \\ &= \inf\left\{\sum_{i=1}^{\infty} F(b_i) - F(a_i) : A \subseteq \bigcup_{i=1}^{\infty}(a_i, b_i]\right\}\end{aligned}$$

The measure μ_F is called the *Lebesgue-Stieltjes measure* associated to F. As we pointed out earlier, our theory not only gives a measure space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu_F)$ with μ_F extending ρ but indeed gives a complete measure space $(\mathbb{R}, \mathcal{M}_{\mu_F}, \mu_F)$ with $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{M}_{\mu_F}$. Here $\mathcal{M}_{\mu_F}$ is the σ-algebra of μ_F^*-measurable sets where μ_F^* is the outer measure derived from ρ as defined in the proof of Theorem 9.

For the rest of this subsection, fix an increasing right-continuous function F. Let $(\mathbb{R}, \mathcal{M}_\mu, \mu)$ be the corresponding complete measure space where μ is the Lebesgue-Stieltjes measure associated to F. We shall next investigate how to "approximate" the measurable sets of this measure space via topologically simple sets. We start by proving a useful lemma.

Lemma 2. *For every $A \in \mathcal{M}_\mu$, we have that*

$$\mu(A) = \inf\left\{\sum_{i=0}^{\infty}\mu((a_i, b_i)) : A \subseteq \bigcup_{i=0}^{\infty}(a_i, b_i)\right\}$$

Proof. Let $A \in \mathcal{M}_\mu$. Let $\{(a_i, b_i)\}_{i\in\mathbb{N}}$ be such that $A \subseteq \bigcup_{i=0}^{\infty}(a_i, b_i)$. Observe that, for every $i \in \mathbb{N}$, we can choose an increasing sequence of numbers $(c_k^i)_{k\in\mathbb{N}}$ such that $c_0^i = a_i$ and $\lim_{k\to\infty} c_k^i = b_i$. Then we have that

$$A \subseteq \bigcup_{i,k\in\mathbb{N}}(c_k^i, c_{k+1}^i]$$

and that

$$\sum_{i=0}^{\infty}\mu((a_i,b_i))=\sum_{i=0}^{\infty}\mu\left(\bigsqcup_{k=0}^{\infty}(c_k^i,c_{k+1}^i]\right)=\sum_{i,k\in\mathbb{N}}F(c_{k+1}^i)-F(c_k^i)=\sum_{i,k\in\mathbb{N}}\mu((c_k^i,c_{k+1}^i])$$

It follows that

$$\mu(A)\leq\inf\left\{\sum_{i=0}^{\infty}\mu((a_i,b_i)):A\subseteq\bigcup_{i=0}^{\infty}(a_i,b_i)\right\}$$

To prove the converse inequality, let $\epsilon>0$ be arbitrary. Then, by the definition of μ, there exist $\{(a_i,b_i]\}_{i\in\mathbb{N}}$ such that $A\subseteq\bigcup_{i=0}^{\infty}(a_i,b_i]$ and $\sum_{i=0}^{\infty}\mu((a_i,b_i])\leq\mu(A)+\epsilon$. Since F is right-continuous, for every $i\in\mathbb{N}$, we can find $\epsilon_i>0$ such that

$$F(b_i+\epsilon_i)-F(b_i)<\frac{\epsilon}{2^{i+1}}$$

On the other hand, $A\subseteq\bigcup_{i=0}^{\infty}(a_i,b_i+\epsilon_i)$ and

$$\begin{aligned}\sum_{i=0}^{\infty}\mu((a_i,b_i+\epsilon_i))\leq\sum_{i=0}^{\infty}\mu((a_i,b_i+\epsilon_i])&=\sum_{i=0}^{\infty}F(b_i+\epsilon_i)-F(a_i)\\&\leq\sum_{i=0}^{\infty}F(b_i)-F(a_i)+\frac{\epsilon}{2^{i+1}}\\&\leq\left(\sum_{i=0}^{\infty}\mu((a_i,b_i])\right)+\epsilon\leq\mu(A)+2\epsilon\end{aligned}$$

Since $\epsilon>0$ was arbitrary, we have that

$$\inf\left\{\sum_{i=0}^{\infty}\mu((a_i,b_i)):A\subseteq\bigcup_{i=0}^{\infty}(a_i,b_i)\right\}\leq\mu(A)$$

which completes the proof. □

As a consequence of this lemma, we have the following.

Theorem 11. *For every $A\in\mathcal{M}_\mu$, we have that*

$$\begin{aligned}\mu(A)&=\inf\{\mu(O):O\supseteq A,\ O\textit{ is open}\}\\&=\sup\{\mu(K):K\subseteq A,\ K\textit{ is compact}\}\end{aligned}$$

Proof. Since any open subset of $\mathbb{R}$ is a countable disjoint union of bounded open intervals, we have that

$$\inf\{\mu(O):O\supseteq A,\ O\text{ is open}\}=\inf\left\{\sum_{i=0}^{\infty}\mu((a_i,b_i)):A\subseteq\bigcup_{i=0}^{\infty}(a_i,b_i)\right\}$$

and hence the first equality follows by Lemma 2. For the second equality, let $A\in\mathcal{M}_\mu$. We split into two cases.

- Assume that A is bounded. If A is also closed, then A is compact and the equality trivially follows. Suppose that A is not closed. For any $\epsilon > 0$, using the first equality, we can find some open set $U \supseteq \overline{A} - A$ such that $\mu(U) \leq \mu(\overline{A} - A) + \epsilon$. Then

$$\begin{aligned}\mu(\overline{A} - U) = \mu(A - U) &= \mu(A) - \mu(A \cap U) \\ &= \mu(A) - (\mu(U) - \mu(U - A)) \\ &\geq \mu(A) - \mu(U) + \mu(\overline{A} - A) \geq \mu(A) - \epsilon\end{aligned}$$

  Since $\epsilon > 0$ was arbitrary and $\overline{A} - U \subseteq A$ is compact, the second equality easily follows.
- Assume that A is unbounded. Let $\epsilon > 0$ be arbitrary. For every $i \in \mathbb{Z}$, using the previous case, we can find a compact set $K_i \subseteq A \cap (i, i+1]$ such that $\mu(K_i) \geq \mu(A \cap (i, i+1]) - \epsilon/2^{|i|}$. Set $H_n = \bigcup_{-n}^{n} K_i$. Then, H_n is compact and $H_n \subseteq A$ for each $n \in \mathbb{N}$ and H_n's are increasing. Consequently,

$$\begin{aligned}\mu(A) \geq \lim_{n\to\infty} \mu(H_n) &= \lim_{n\to\infty} \sum_{i=-n}^{n} \mu(K_i) \\ &\geq \lim_{n\to\infty} \sum_{i=-n}^{n} \mu(A \cap (i, i+1]) - \epsilon/2^{|i|} \\ &\geq \lim_{n\to\infty} \mu\left(\bigsqcup_{i=-n}^{n} A \cap (i, i+1]\right) - 3\epsilon \\ &\geq \mu(A) - 3\epsilon\end{aligned}$$

  As $\epsilon > 0$ is arbitrary, the second equality follows.

□

As a corollary of Theorem 11, we have the following characterization of measurable sets in $\mathcal{M}_\mu$.

Corollary 12. *Let $A \subseteq \mathbb{R}$. The following are equivalent.*

(a) $A \in \mathcal{M}_\mu$.
(b) $A = F \cup N_1$ *for some* $F, N_1 \subseteq \mathbb{R}$ *where* F *is in* $\mathbf{\Sigma}^0_2$ *and* N_1 *is a subset of some* μ*-null set in* $\mathcal{B}(\mathbb{R})$.
(c) $A = G - N_2$ *for some* $G, N_2 \subseteq \mathbb{R}$ *where* G *is in* $\mathbf{\Pi}^0_2$ *and* N_2 *is a subset of some* μ*-null set in* $\mathcal{B}(\mathbb{R})$.

Proof. (b) and (c) separately imply (a) since $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{M}_\mu$ and $(\mathbb{R}, \mathcal{M}_\mu, \mu)$ is complete. To show that (a) implies (c), assume that $A \in \mathcal{M}_\mu$. Observe that $(\mathbb{R}, \mathcal{M}_\mu, \mu)$ is σ-finite and hence we can write $A = \bigcup_{k\in\mathbb{N}} A_k$ for some *disjoint* sets $A_k \in \mathcal{M}_\mu$ with $\mu(A_k) < \infty$. By Theorem 11, we can find a sequence of open sets $O^k_n \supseteq A_k$ such that

$$\mu(O^k_n) - \frac{1}{(n+1)2^{k+1}} \leq \mu(A_k) \leq \mu(O^k_n)$$

Set $G = \bigcap_{n\in\mathbb{N}} \bigcup_{k\in\mathbb{N}} O_n^k$. Then clearly $A \subseteq G$ and G is in $\mathbf{\Pi}_2^0$. Moreover, for every $n \in \mathbb{N}$, we have

$$\begin{aligned}\mu(G-A) \leq \mu\left(\left(\bigcup_{k\in\mathbb{N}} O_n^k\right) - A\right) &\leq \mu\left(\bigcup_{k\in\mathbb{N}} (O_n^k - A)\right) \\ &\leq \mu\left(\bigcup_{k\in\mathbb{N}} (O_n^k - A_k)\right) \\ &\leq \sum_{k\in\mathbb{N}} \mu(O_n^k - A_k) \\ &= \sum_{k\in\mathbb{N}} \mu(O_n^k) - \mu(A_k) \leq \frac{1}{n+1}\end{aligned}$$

It follows that $\mu(G-A) = 0$. We now show that $G - A$ is a subset of a Borel μ-null set. Using Theorem 11, for every $n \in \mathbb{N}$, one can find a $\mathbf{\Pi}_2^0$-set $H_n \subseteq \mathbb{R}$ such that $\mu(H_n) = 0$ and $((G-A) \cap [-n,n]) \subseteq H_n$. Setting $H = \bigcup_{n\in\mathbb{N}} H_n$, we have that H is in $\mathbf{\Sigma}_3^0$ and $G - A \subseteq H$ and $\mu(H) = 0$. The statement (c) now follows as $A = G - (G - A)$.

Having shown the equivalence of (a) and (c), we now show that (a) implies (b). Assume that $A \in \mathcal{M}_\mu$. Then $A^c \in \mathcal{M}_\mu$. Now, since (a) implies (c), $A^c = G - N_2$ for some $G, N_2 \subseteq \mathbb{R}$ where G is in $\mathbf{\Pi}_2^0$ and N_2 is a subset of some μ-null set in $\mathcal{B}(\mathbb{R})$. But then, G^c is in $\mathbf{\Sigma}_2^0$ and $A = (A^c)^c = (G - N_2)^c = G^c \cup N_2$, which is what we wanted to show. This completes the proof. $\square$

We would like to remark that Corollary 12 indeed shows that $(\mathbb{R}, \mathcal{M}_\mu, \mu)$ is the completion of the measure space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$.

2.5. **The Lebesgue measure.** We now turn our attention to the most important measure on $\mathbb{R}$, namely, the Lebesgue-Stieltjes measure corresponding to the identity function $F(x) = x$. By the machinery that we have developed so far, there exists a complete measure space

$$(\mathbb{R}, \mathfrak{L}, \mathbf{m})$$

such that $\mathcal{B}(\mathbb{R}) \subseteq \mathfrak{L}$ and $\mathbf{m}((a,b]) = b - a$ for every $a < b$ and

$$\begin{aligned}\mathbf{m}(A) &= \inf\left\{\sum_{i=1}^{\infty} \mathbf{m}((a_i,b_i]) : A \subseteq \bigcup_{i=1}^{\infty} (a_i,b_i]\right\} \\ &= \inf\left\{\sum_{i=1}^{\infty} b_i - a_i : A \subseteq \bigcup_{i=1}^{\infty} (a_i,b_i]\right\}\end{aligned}$$

for every $A \in \mathfrak{L}$. Moreover, since F is also left-continuous, one can show that $\mathbf{m}((a,b)) = \mathbf{m}((a,b]) = b - a$ and hence, by Theorem 2, we have

$$\mathbf{m}(A) = \inf\left\{\sum_{i=1}^{\infty} b_i - a_i : A \subseteq \bigcup_{i=1}^{\infty} (a_i,b_i)\right\}$$

The measure $\mathbf{m}$ is called the *Lebesgue measure* on $\mathbb{R}$, the σ-algebra is called the *Lebesgue σ-algebra* and the elements of $\mathfrak{L}$ are called the *Lebesgue measurable sets* of $\mathbb{R}$. By the construction in the proof of Carathéodory's extension theorem, we have that a set $A \in \mathfrak{L}$ if and only if A is $\mathbf{m}^*$-measurable, i.e.

$$\mathbf{m}^*(X) = \mathbf{m}^*(X \cap A) + \mathbf{m}^*(X \cap A^c) \text{ for all } X \subseteq \mathbb{R}$$

where $\mathbf{m}^*$ denotes the Lebesgue outer measure given by the same formula defining $\mathbf{m}$. By Theorem 12, we also have another characterization of Lebesgue measurable subsets of $\mathbb{R}$, namely, that $A \in \mathfrak{L}$ if and only if $A = B \cup H$ for some $B \in \mathbf{\Sigma}^0_2$ and for some $H \subseteq N$ where $N \in \mathcal{B}(\mathbb{R})$ is $\mathbf{m}$-null. We now show that the Lebesgue measure is invariant under translations and dilations.

Theorem 13. *For any $c \geq 0$ and $A \in \mathfrak{L}$, we have that $cA,\ c + A \in \mathfrak{L}$. Moreover, $\mathbf{m}(cA) = c \cdot \mathbf{m}(A)$ and $\mathbf{m}(c + A) = \mathbf{m}(A)$.*

Proof. Let $c \geq 0$ and $A \in \mathfrak{L}$. The claims are trivial if $c = 0$. Assume that $c > 0$. Observe that if we translate by c (respectively, dilate by c) the elements of a covering of a set $S \subseteq \mathbb{R}$ via right-closed half-open intervals, then we get a covering of $c + S$ (respectively, of cS) via right-closed half-open intervals. Conversely, coverings of $c+S$ and cS via right-closed half-open intervals will canonically induce coverings of S via right-closed half-open intervals. With this in mind, it is difficult not to prove that $\mathbf{m}^*(c+S) = \mathbf{m}^*(S)$ and $\mathbf{m}^*(cS) = c \cdot \mathbf{m}^*(S)$ for all $S \subseteq \mathbb{R}$. Since $\mathbf{m} = \mathbf{m}^* \restriction \mathfrak{L}$, it now remains to show that $c + A \in \mathfrak{L}$ and $cA \in \mathfrak{L}$.

As $A \in \mathfrak{L}$, we can write A as $A = B \cup H$ for some $B \in \mathcal{B}(\mathbb{R})$ and some $H \subseteq N$ where $N \in \mathcal{B}(\mathbb{R})$ is a $\mathbf{m}$-null set. On the other hand, $c+B$ and $c+N$ are easily seen to be Borel[4], $\mathbf{m}(c + N) = \mathbf{m}(N) = 0$ by our previous observation, $c + H \subseteq c + N$ and $c+A = (c+B) \cup (c+H)$. It now follows from $\mathcal{B}(\mathbb{R}) \subseteq \mathfrak{L}$ and the completeness of $(\mathbb{R}, \mathfrak{L}, \mathbf{m})$ that $c+A \in \mathfrak{L}$. One can show via a similar argument that $cA \in \mathfrak{L}$. ☐

The Lebesgue measure also satisfies that $\mathbf{m}(A) = \mathbf{m}(-A)$ and that $-A \in \mathfrak{L}$ if and only if $A \in \mathfrak{L}$. Together with Theorem 13, this observation implies that $\mathbf{m}(cA) = |c| \cdot \mathbf{m}(A)$ and $\mathbf{m}(c + A) = \mathbf{m}(A)$ for all $c \in \mathbb{R}$. The reader is expected to prove this generalization if he or she has not already. An important consequence of the translation invariance of the Lebesgue measure is the following.

Theorem 14. *There exists non-Lebesgue measurable sets, that is, $\mathfrak{L} \neq \mathcal{P}(\mathbb{R})$.*

Proof. This immediately follows from Theorem 1 and Theorem 13 since m is a measure. ☐

---

[4]To see this, show that $C = \{S \subseteq \mathbb{R} : c + S \in \mathcal{B}(\mathbb{R})\}$ is a σ-algebra containing open sets and hence $\mathcal{B}(\mathbb{R}) \subseteq C$.

Observe that a set $A \subseteq \mathbb{R}$ is of Lebesgue measure zero, that is,

$$\mathbf{m}(A) = \inf \left\{ \sum_{i=1}^{\infty} b_i - a_i : A \subseteq \bigcup_{i=1}^{\infty} (a_i, b_i) \right\} = 0$$

if and only if for every $\epsilon > 0$ there exists a sequence $(I_k)_{k\in\mathbb{N}}$ of open intervals such that $A \subseteq \bigcup_{k\in\mathbb{N}} I_k$ and $\sum_{k=0}^{\infty} |I_k| < \epsilon$, where $|I_k|$ denotes the length of this interval. From a measure-theoretic point of view, sets of Lebesgue measure zero are "negligible" subsets of $\mathbb{R}$. For this reason, it is natural to ask the following question. What sets are of Lebesgue measure zero?

Proposition 3. *Every countable subset of* $\mathbb{R}$ *has Lebesgue measure zero.*

Proof. Let $A \subseteq \mathbb{R}$ be countable. Say, $A = \{a_n\}_{n\in\mathbb{N}}$. Let $\epsilon > 0$. Then we have that

$$A \subseteq \bigcup_{n\in\mathbb{N}} B\left(a_n, \frac{\epsilon}{2^{n+3}}\right)$$

and $\sum_{n=0}^{\infty} |B(a_n, \frac{\epsilon}{2^{n+3}})| = \frac{\epsilon}{2} < \epsilon$. This completes the proof. □

The next natural question is the following. Are there uncountable subsets of the real line that are of Lebesgue measure zero? This brings us to the Cantor set, which the author believes is one of the most profound instructional examples and which is often used to construct counter examples to various reasonable but false conjectures.

2.6. **The Cantor set.** Consider the set $2^{\mathbb{N}} = \{(a_i)_{i\in\mathbb{N}} : \forall i \in \mathbb{N}\ a_i \in \{0,1\}\}$ consisting of infinite binary sequences equipped with the topology τ induced by the metric

$$d((a_i)_{i\in\mathbb{N}}, (b_i)_{i\in\mathbb{N}}) = \begin{cases} 0 & \text{if } (a_i)_{i\in\mathbb{N}} = (b_i)_{i\in\mathbb{N}} \\ 2^{-\min\{i\in\mathbb{N}:\ a_i \neq b_i\}} & \text{if } (a_i)_{i\in\mathbb{N}} \neq (b_i)_{i\in\mathbb{N}} \end{cases}$$

It is readily verified that $(2^{\mathbb{N}}, \tau)$ is a compact totally-disconnected perfect Polish space, known as the *Cantor space*. Consider the function $f : 2^{\mathbb{N}} \to \mathbb{R}$ given by

$$f((a_i)_{i\in\mathbb{N}}) = \sum_{i=0}^{\infty} \frac{2a_i}{3^{i+1}}$$

In other words, $f((a_i)_{i\in\mathbb{N}})$ is the number in $[0,1]$ with the ternary expansion $(0, (2a_0)(2a_1)(2a_2)\dots)_3$. It is straightforward to check that f is continuous. We now check that f is an injection. Assume towards a contradiction that for distinct sequences $(a_i)_{i\in\mathbb{N}}$ and $(b_i)_{i\in\mathbb{N}}$ we have $f((a_i)_{i\in\mathbb{N}}) = f((b_i)_{i\in\mathbb{N}})$. Then we get

$$0 = \sum_{i=0}^{\infty} \frac{2(a_i - b_i)}{3^{i+1}}$$

Let $k \in \mathbb{N}$ be least such that $a_k \neq b_k$. Without loss of generality, we may assume that $a_k = 1$ and $b_k = 0$. Then clearly

$$0 = \sum_{i=0}^{\infty} \frac{2(a_i - b_i)}{3^{i+1}} = \frac{2}{3^{k+1}} + \sum_{i=k+1}^{\infty} \frac{2(a_i - b_i)}{3^{i+1}}$$

On the other hand, it follows from triangle inequality that

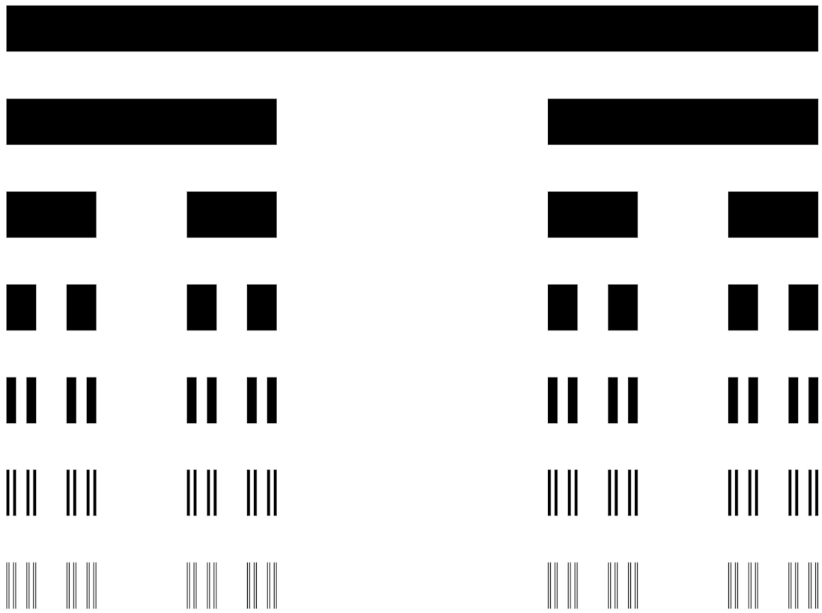
$$\frac{2}{3^{k+1}} = \left| \sum_{i=k+1}^{\infty} \frac{2(a_i - b_i)}{3^{i+1}} \right| \leq \sum_{i=k+1}^{\infty} \left| \frac{2(a_i - b_i)}{3^{i+1}} \right| \leq \sum_{i=k+1}^{\infty} \left| \frac{2}{3^{i+1}} \right| = \frac{1}{3^{k+1}}$$

which is a contradiction. Therefore, f is an injection. Since f is a continuous injection and $|2^{\mathbb{N}}| = |\mathbb{R}| = \mathfrak{c}$, its image is a compact perfect subset of $\mathbb{R}$ with cardinality $\mathfrak{c}$. Its image $\mathcal{C} = f[2^{\mathbb{N}}]$ is called the *Cantor set*.

In other words, the Cantor set $\mathcal{C}$ consists of the numbers in $[0, 1]$ which has some ternary expansion that does not contain the digit 1. Another way to describe $\mathcal{C}$ is the following recursive procedure:

- Start with the interval $[0, 1]$.
- At every stage, for each closed interval, delete the *open* middle one-third.
- Repeat the second step for countable many stages.
- $\mathcal{C}$ is the intersection of all sets obtained after these stages.

The union of open middle one-thirds deleted at the stage $k \geq 1$ is the set of numbers in $[0, 1]$ which have ternary expansions containing 0 or 2 at their i-th digits for $1 \leq i < k$ and containing 1 at their k-th digits that is not followed by all 2's or all 0's. This recursive procedure can be visualized as follows.



Being a compact set, $\mathcal{C}$ is closed and hence is Lebesgue measurable. We now compute its Lebesgue measure. Set

$$D_k = \left\{ x \in [0,1] : x = \sum_{i=0}^{\infty} \frac{a_i}{3^{i+1}},\ a_i \in \{0,1,2\} \text{ and } a_i \neq 1 \text{ for } 0 \leq i < k \text{ and } a_k = 1 \right\}$$

for each $k \in \mathbb{N}$. In other words, D_k is the set of numbers in $[0,1]$ which have some ternary expansion that contains 1 at its $(k+1)$-st digit and not contains 1 at its earlier digits. We would like to remark that the set D_k is the closure of the set removed at stage $k+1$ in the previous recursive construction. Note that the sets D_k are disjoint and

$$[0,1] = \mathcal{C} \cup \bigsqcup_{k=0}^{\infty} D_k$$

Moreover, the sets $\mathcal{C} \cap D_k$ are at most countable since there are only countably many numbers in $[0,1]$ which has multiple (indeed, two) ternary expansions.[5] Consequently, we have that $\mathbf{m}(D_k - \mathcal{C}) = \mathbf{m}(D_k)$. It is also not difficult to see that $\mathbf{m}(D_k) = 2^k/3^{k+1}$. It follows that

$$\mathbf{m}(\mathcal{C}) = 1 - \mathbf{m}\left(\bigsqcup_{k=0}^{\infty} (D_k - \mathcal{C})\right) = 1 - \sum_{k=0}^{\infty} \mathbf{m}(D_k - \mathcal{C}) = 1 - \sum_{k=0}^{\infty} \mathbf{m}(D_k) = 1 - \sum_{k=0}^{\infty} \frac{2^k}{3^{k+1}} = 0$$

[5]With some effort, the reader can verify that $\mathcal{C} \cap D_k$ has indeed 2^{k+1} elements.

Therefore, $\mathcal{C}$ is of Lebesgue measure zero. Having shown the existence of a set of Lebesgue measure zero with cardinality $\mathfrak{c}$, we have the following corollaries.

Corollary 15. $|\mathfrak{L}| = |\mathcal{P}(\mathbb{R})|$.

Proof. Since $\mathfrak{L} \subseteq \mathcal{P}(\mathbb{R})$, we have that $|\mathfrak{L}| \leq |\mathcal{P}(\mathbb{R})|$. For the converse inequality, observe that, since $(\mathbb{R}, \mathfrak{L}, \mathbf{m})$ is complete and $\mathbf{m}(\mathcal{C}) = 0$, we have $\mathcal{P}(\mathcal{C}) \subseteq \mathfrak{L}$ and hence, that $|\mathcal{C}| = \mathfrak{c} = |\mathbb{R}|$ implies that $|\mathcal{P}(\mathcal{C})| = |\mathcal{P}(\mathbb{R})| \leq |\mathfrak{L}|$. □

Corollary 16. *There are non-Borel Lebesgue measurable sets, that is,* $\mathcal{B}(\mathbb{R}) \neq \mathfrak{L}$.

Proof. It follows from Theorem 3, Corollary 15 and Cantor's theorem on the cardinality of power sets that $|\mathcal{B}(\mathbb{R})| = \mathfrak{c} = |\mathbb{R}| < |\mathcal{P}(\mathbb{R})| = |\mathfrak{L}|$. Thus $\mathcal{B}(\mathbb{R}) \neq \mathfrak{L}$. □

Before we conclude this subsection, we would like to remark that, in the recursive construction of the Cantor set, if one removes the open middle one-n-th instead of the open middle one-third, where $n > 3$, then the corresponding construction would end up giving a compact perfect totally-disconnected subset of $\mathbb{R}$ with positive measure. Such sets are usually called *fat Cantor sets* or *Smith-Volterra-Cantor sets*. The curious reader may Google these terms to see examples of such constructions.

2.7. **Stairway to hell.** We shall next construct a function based on the Cantor set which has some amazing properties. Define the function $F : [0,1] \to [0,1]$ by

$$F(x) = \begin{cases} \displaystyle\sum_{i=0}^{\infty} \frac{a_i}{2^{i+1}} & \text{if } x \in \mathcal{C} \text{ and } x = \displaystyle\sum_{i=0}^{\infty} \frac{2a_i}{3^{i+1}} \\ \sup\{F(t) : t \leq x,\ t \in \mathcal{C}\} & \text{if } x \notin \mathcal{C} \end{cases}$$

This function is well-defined because any element of $\mathcal{C}$ has a unique ternary expansion not containing the digit 1. We will now show that F is increasing. Let $x, y \in [0,1]$. Assume that $x \leq y$. We have the following four cases.

- If $x \notin \mathcal{C}$ and $y \notin \mathcal{C}$, then $F(x) = F(y)$ since
$$\{F(t) : t \leq x,\ t \in \mathcal{C}\} \subseteq \{F(t) : t \leq y,\ t \in \mathcal{C}\}$$
- If $x \in \mathcal{C}$ and $y \notin \mathcal{C}$, then $F(x) \leq F(y)$ since $F(x) \in \{F(t) : t \leq y,\ t \in \mathcal{C}\}$.
- If $x \in \mathcal{C}$ and $y \in \mathcal{C}$, then $F(x) \leq F(y)$ since each ternary digit of y is greater than or equal to the corresponding ternary digit of x.
- If $x \notin \mathcal{C}$ and $y \in \mathcal{C}$, then $F(x) \leq F(y)$ since $F(t) \leq F(y)$ for every $t \leq x \leq y$ with $t \in \mathcal{C}$, by the previous case.

Therefore F is increasing. Recall that every number in $[0,1]$ has a binary expansion, i.e. is of the form $\sum_{i=0}^{\infty} \frac{a_i}{2^{i+1}}$. It follows that F is surjective. In fact, we have that $F[\mathcal{C}] = [0,1]$. We shall next prove the continuity of F.

Theorem 17. *F is continuous on* $[0,1]$.

Proof. We first show that F is continuous outside the Cantor set. Let $x \in [0,1]-\mathcal{C}$. Since $[0,1]-\mathcal{C}$ is open, there exists $\delta > 0$ such that $B(x,\delta) \subseteq [0,1]-\mathcal{C}$. But then, by the definition of F, we have that $F(y)=F(x)$ whenever $y \in B(x,\delta)$. Therefore, $F'(x)=0$ and F is continuous at x.

We now show that F is continuous at every point in $\mathcal{C}$. Let $x \in \mathcal{C}$ and $\epsilon > 0$. Since $x \in \mathcal{C}$, there exists a unique sequence $(a_n)_{n\in\mathbb{N}} \in 2^{\mathbb{N}}$ such that $x = \sum_{i=0}^{\infty} \frac{2a_i}{3^{i+1}}$. We can also find some integer $k \geq 2$ such that $1/2^k < \epsilon$. We split into several cases.

- Assume that $(a_n)_{n\in\mathbb{N}} \in 2^{\mathbb{N}}$ is not eventually 0 or eventually 1, i.e. the digits in the ternary expansion of x is not eventually 0 or eventually 2. Choose

$$\delta = \min\left\{ x - \sum_{i=0}^{k-1} \frac{2a_i}{3^{i+1}}, \left(\sum_{i=0}^{k-1} \frac{2a_i}{3^{i+1}} + \sum_{i=k}^{\infty} \frac{2}{3^{i+1}}\right) - x \right\}$$

Since $(a_n)_{n\in\mathbb{N}} \in 2^{\mathbb{N}}$ is not eventually 0 or eventually 1, we have that $\delta > 0$. Assume that $|x-y| < \delta$. If we have $x-\delta < y \leq x$, then, F being increasing implies

$$F(x)-F(y) \leq F(x) - F\left(\sum_{i=0}^{k-1} \frac{2a_i}{3^{i+1}}\right) = \sum_{i=0}^{\infty} \frac{a_i}{2^{i+1}} - \sum_{i=0}^{k-1} \frac{a_i}{2^{i+1}} \leq \sum_{i=k}^{\infty} \frac{1}{2^{i+1}} = \frac{1}{2^k} < \epsilon$$

Similarly, if we have $x \leq y < x+\delta$, then

$$F(y)-F(x) \leq F\left(\sum_{i=0}^{k-1} \frac{2a_i}{3^{i+1}} + \sum_{i=k}^{\infty} \frac{2}{3^{i+1}}\right) - F(x) = \sum_{i=k}^{\infty} \frac{1-a_i}{2^{i+1}} \leq \frac{1}{2^k} < \epsilon$$

This shows that F is continuous at x.

- Assume that $(a_n)_{n\in\mathbb{N}} \in 2^{\mathbb{N}}$ is eventually 0, i.e. the digits in the ternary expansion of x is eventually 0. Set

$$\delta = \left(\sum_{i=0}^{k-1} \frac{2a_i}{3^{i+1}} + \sum_{i=k}^{\infty} \frac{2}{3^{i+1}}\right) - x$$

Then $\delta > 0$. If we have $x \leq y < x+\delta$, then, as before,

$$F(y)-F(x) \leq F\left(\sum_{i=0}^{k-1} \frac{2a_i}{3^{i+1}} + \sum_{i=k}^{\infty} \frac{2}{3^{i+1}}\right) - F(x) = \sum_{i=k}^{\infty} \frac{1-a_i}{2^{i+1}} = \frac{1}{2^k} < \epsilon$$

which shows that F is right-continuous at x. We will now show that left-continuous at x whenever $x \neq 0$. Assume that $x \neq 0$, say, $x = \sum_{i=0}^{k} \frac{2a_i}{3^{i+1}}$ with $a_k \neq 0$. Set $\delta = 1/3^{k+1}$. If we have $x-\delta < y < x$, then any ternary expansion of y contains 1 at its $(k+1)$-st digit and hence $y \notin \mathcal{C}$. Moreover,

$$x - \delta = \sum_{i=0}^{k-1} \frac{2a_i}{3^{i+1}} + \frac{1}{3^{k+1}} = \sum_{i=0}^{k-1} \frac{2a_i}{3^{i+1}} + \sum_{i=k+1}^{\infty} \frac{2}{3^{i+1}} \in \mathcal{C}$$

Consequently, as F is increasing, it follows from the definition of F that $F(x-\delta) = F(y)$ for any $x - \delta < y < x$. On the other hand, we have

$$F(x) = \sum_{i=0}^{k} \frac{a_i}{2^{i+1}} = \sum_{i=0}^{k-1} \frac{a_i}{2^{i+1}} + \frac{1}{2^{k+1}} = \sum_{i=0}^{k-1} \frac{a_i}{2^{i+1}} + \sum_{i=k+1}^{\infty} \frac{1}{2^{i+1}} = F(x-\delta)$$

Therefore, $F(x) = F(y)$ for any $x - \delta < y < x$. This shows that F is left-continuous at x whenever $x \neq 0$. Therefore, F is continuous at x.

- Assume that $(a_n)_{n\in\mathbb{N}} \in 2^{\mathbb{N}}$ is eventually 2, i.e. the digits in the ternary expansion of x is eventually 2. Set

$$\delta = x - \sum_{i=0}^{k-1} \frac{2a_i}{3^{i+1}}$$

Then $\delta > 0$. If we have $x - \delta < y \leq x$, then, as before,

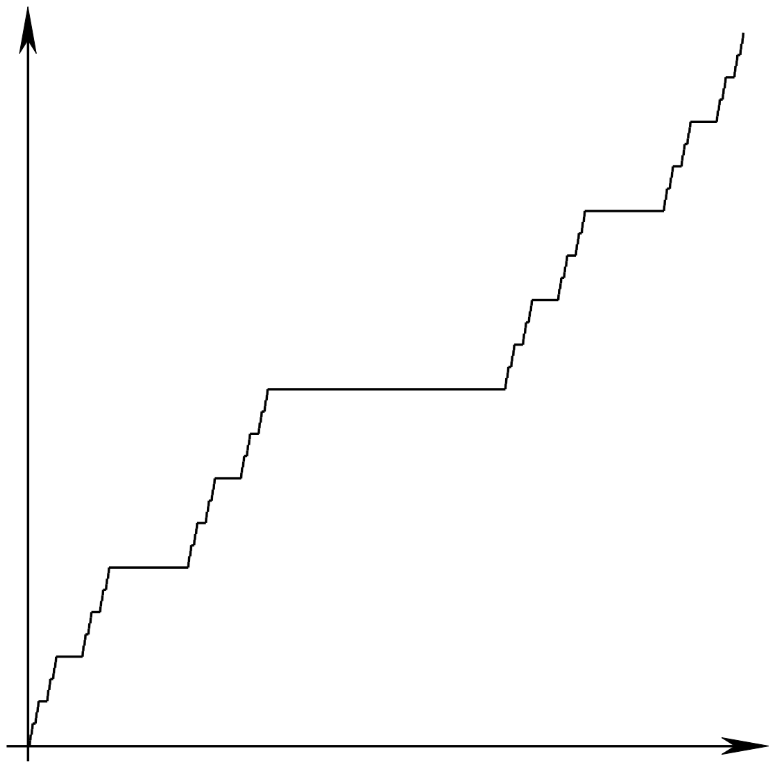
$$F(x) - F(y) \leq F(x) - F\left(\sum_{i=0}^{k-1} \frac{2a_i}{3^{i+1}}\right) = \sum_{i=0}^{\infty} \frac{a_i}{2^{i+1}} - \sum_{i=0}^{k-1} \frac{a_i}{2^{i+1}} \leq \sum_{i=k}^{\infty} \frac{1}{2^{i+1}} = \frac{1}{2^k} < \epsilon$$

which shows that F is left-continuous at x. We will now show that F is right-continuous at x whenever $x \neq 1$. Assume that $x \neq 1$, say, we have $x = \sum_{i=0}^{k} \frac{2a_i}{3^{i+1}} + \sum_{i=k+1}^{\infty} \frac{2}{3^{i+1}}$ with $a_k \neq 2$. Set $\delta = 1/3^{k+1}$. If we have $x < y < x + \delta$, then any ternary expansion of y contains 1 at its $(k+1)$-st digit and hence $y \notin \mathcal{C}$. Consequently, as F is increasing, it follows from the definition of F that $F(x) = F(y)$ for any $x < y < x + \delta$. This shows that F is right-continuous at x whenever $x \neq 1$.

This completes the proof that F is continuous. □

The function F is called the *Cantor function*. As we have shown, it is an increasing surjective continuous function from the closed unit interval to itself whose derivative is zero almost everywhere. It is sometimes called the *Devil's staircase* since its graph (approximately) looks as follows.[6]

[6]Imagine trying to climb these stairs from $(0,0)$ to $(1,1)$. You are *somehow* climbing up without moving vertically except on a "negligible" set.



Being a continuous function on a compact set, the Cantor function is uniformly continuous but is not absolutely continuous.

2.8. **Strong measure zero sets.** In this subsection, we shall introduce a strengthening of the notion of being of measure zero. Recall that a set $A \subseteq \mathbb{R}$ is of (Lebesgue) measure zero if and only if A can be covered with countably many open intervals whose total length can be made arbitrarily small. By demanding to control the length of *each* of these open intervals, we obtain the following stronger notion.

A set $A \subseteq \mathbb{R}$ is *of strong measure zero* if and only if for every sequence $(\epsilon_n)_{n\in\mathbb{N}}$ of positive real numbers, there exists a sequence $(I_n)_{n\in\mathbb{N}}$ of bounded open intervals such that $A \subseteq \bigcup_{n=0}^{\infty} I_n$ and $|I_n| < \epsilon_n$ for all $n \in \mathbb{N}$. It is straightforward to check that every strong measure zero set is of measure zero. However, the converse statement does not hold.

Proposition 4. *The Cantor set $\mathcal{C}$ is not of strong measure zero.*

Proof. Assume towards a contradiction that $\mathcal{C}$ is of strong measure zero. Given $\epsilon_n = 1/3^{n+1}$ for every $n \in \mathbb{N}$, by assumption, we can find open intervals $(I_n)_{n\in\mathbb{N}}$ such that $|I_n| < 1/3^{n+1}$ and $\mathcal{C} = \bigcup_{n\in\mathbb{N}} I_n$.

Observe that I_0 does not intersect at least one of $[0, 1/3]$ and $[2/3, 1]$ since it has length $1/3$. Say, it does not intersect $[a_0/3, (a_0 + 1)/3]$. Similarly, I_1 does not intersect at least one of $[a_0/9, (a_0 + 1)/9]$ or $[(a_0 + 2)/9, (a_0 + 3)/9]$ since it has length $1/9$. Say, it does not intersect $[a_1/9, (a_1 + 1)/9]$. Continuing inductively in this manner, we can find a sequence of nested closed intervals $(J_n)_{n\in\mathbb{N}}$ such that I_n does not intersect J_n. But, by compactness, we have that $\bigcap_{n\in\mathbb{N}} J_n \neq \emptyset$. Let $x \in \bigcap_{n\in\mathbb{N}} J_n$. By the recursive construction of $\mathcal{C}$, we have that $x \in \mathcal{C}$ but $x \notin \bigcup_{n\in\mathbb{N}} I_n$ as J_n does not intersect I_n, which is a contradiction. □

Which sets are of strong measure zero? Imitating the proof of Proposition 3, one can easily see that countable sets are of strong measure zero, as follows. Given $A = \{a_n : n \in \mathbb{N}\}$ and $(\epsilon_n)_{n\in\mathbb{N}}$, we can set $I_n = B(a_n, \epsilon_n/4)$ in which case $|I_n| < \epsilon_n$ and $A \subseteq \bigcup_{n\in\mathbb{N}} I_n$. The next obvious question is the following. Does the converse statement hold?

The Borel conjecture is the statement that every strong measure zero set is countable. Sierpiński proved in [Sie28] that, assuming the Continuum Hypothesis[7] in addition to ZFC, there are uncountable strong measure zero sets. Laver proved in [Lav76] that if ZFC is consistent, then so is ZFC+"the Borel conjecture holds". Since the Continuum Hypothesis is relatively consistent with ZFC, combining these results, we have the following.

Theorem 18. *If the axioms of ZFC are consistent, then the Borel conjecture cannot be proven or disproven using the axioms of ZFC.*

This theorem is one of the many connections between abstract measure theory and advanced set theory.

2.9. **Exercises.** Below you shall find some exercises that you can work on regarding the topics in this section. These exercises are *not* to be handed in as homework assignments.

- Exercises 10, 11, 18, 19, 24, 30, 33 from Chapter 1 of [Fol99].
- Exercises 2,3 from Chapter 1.2, Exercise 9 from Chapter 1.3 and Exercise 7 from Chapter 1.4 of [Coh93].
- Show that for every continuous map $f : (0, 1) \to \mathbb{R}$, there exists a continuous map $g : (0, 1) \to \mathbb{R}$ such that $range(f) = range(g)$ and $g'(x) = 0$ almost everywhere, where $(0, 1)$ is endowed with its Lebesgue measure.

3. Functions

3.1. **Measurable functions.** In this subsection, we shall introduce and investigate measurable functions which are the "morphisms" between measurable spaces.

[7]Recall that the Continuum hypothesis is the statement that there are not sets A such that $|\mathbb{N}| < |A| < \mathbb{R}$.

Let $(\mathbf{X}, \mathcal{M})$ and $(\mathbf{Y}, \mathcal{N})$ be measurable spaces. A function $f : \mathbf{X} \to \mathbf{Y}$ is said to be *measurable* if $f^{-1}[E] \in \mathcal{M}$ for every $E \in \mathcal{N}$.[1] In other words, a map between measurable spaces is called measurable if the inverse images of measurable sets are measurable. It is immediately seen from the definition that the composition of measurable maps is measurable, whenever such a composition is possible.

In what follows, we shall often be dealing with functions of the form $f : \mathbf{X} \to \mathbb{R}$. Let us introduce some terminology regarding some of these functions. In the case that $\mathbf{X}$ is a topological space, a function $f : \mathbf{X} \to \mathbb{R}$ will be called *Borel measurable* (or simply, *Borel*) if it is $(\mathcal{B}(\mathbf{X}), \mathcal{B}(\mathbb{R}))$-measurable. In the special case that $(\mathbf{X}, \mathcal{M})$ is equal to $(\mathbb{R}, \mathfrak{L})$, a function $f : \mathbb{R} \to \mathbb{R}$ will be called *Lebesgue measurable* if it is $(\mathfrak{L}, \mathcal{B}(\mathbb{R}))$-measurable.

It turns out that, in order to guarantee measurability, it suffices to check the inverse images of generators.

Proposition 5. *Let $(\mathbf{X}, \mathcal{M})$ and $(\mathbf{Y}, \mathcal{N})$ be measurable spaces. Let $\mathcal{E} \subseteq \mathcal{P}(\mathbf{Y})$ be such that $\mathcal{M}(\mathcal{E}) = \mathcal{N}$. Then a map $f : \mathbf{X} \to \mathbf{Y}$ is measurable if and only if $f^{-1}[E] \in \mathcal{M}$ for every $E \in \mathcal{E}$.*

Proof. The left-to-right direction is trivial since $\mathcal{E} \subseteq \mathcal{N}$. To prove the right-to-left direction, assume that $f^{-1}[E] \in \mathcal{M}$ for every $E \in \mathcal{E}$. Set

$$\Omega = \{A \subseteq \mathbf{Y} : \ f^{-1}[A] \in \mathcal{M}\}$$

By assumption, we have that $\mathcal{E} \subseteq \Omega$. It suffices to show that $\mathcal{N} = \mathcal{M}(\mathcal{E}) \subseteq \Omega$, which would follow from that Ω is a σ-algebra. Let $A \in \Omega$ and $A_1, A_2, \dots \in \Omega$. Then $f^{-1}[A] \in \mathcal{M}$ and $f^{-1}[A_i]$ for every $i \in \mathbb{N}^+$. Since $\mathcal{M}$ is a σ-algebra, we have that $f^{-1}[A^c] = \left(f^{-1}[A]\right)^c \in \mathcal{M}$ and

$$f^{-1}\left[\bigcup_{i=1}^{\infty} A_i\right] = \bigcup_{i=1}^{\infty} f^{-1}[A_i] \in \mathcal{M}$$

Therefore, $A^c \in \Omega$ and $\bigcup_{i=1}^{\infty} A_i \in \Omega$. This shows that Ω is a σ-algebra. □

Since the Borel σ-algebra of a topological space is generated by its open sets and the inverse images of open sets are open under continuous functions, as a corollary to Proposition 5, we have that any continuous function $f : \mathbf{X} \to \mathbf{Y}$ is automatically $(\mathcal{B}(\mathbf{X}), \mathcal{B}(\mathbf{Y}))$-measurable. Combining Proposition 5 and Proposition 2, we also have the following characterization of measurable functions with codomain $\mathbb{R}$.

Proposition 6. *Let $(\mathbf{X}, \mathcal{M})$ be a measurable space and $f : \mathbf{X} \to \mathbb{R}$. Then the following are equivalent.*

- *f is measurable.*

[1] To be more precise, what we should really say is that f is $(\mathcal{M}, \mathcal{N})$-*measurable*. However, the endowed σ-algebras are often understood from the context and so, we shall drop the prefix unless it is necessary.

- $f^{-1}[\ (a,b)\ ] \in \mathcal{M}$ *for all* $a, b \in \mathbb{R}$.
- $f^{-1}[\ (a,b)\ ] \in \mathcal{M}$ *for all* $a, b \in \mathbb{Q}$.
- $f^{-1}[\ [a,b]\ ] \in \mathcal{M}$ *for all* $a, b \in \mathbb{R}$.
- $f^{-1}[\ [a,b)\ ] \in \mathcal{M}$ *for all* $a, b \in \mathbb{R}$.
- $f^{-1}[\ (a,\infty)\ ] \in \mathcal{M}$ *for all* $a, b \in \mathbb{R}$.
- $f^{-1}[\ (-\infty,a)\ ] \in \mathcal{M}$ *for all* $a, b \in \mathbb{R}$.
- $f^{-1}[\ [a,\infty)\ ] \in \mathcal{M}$ *for all* $a, b \in \mathbb{R}$.
- $f^{-1}[\ (-\infty,a]\ ] \in \mathcal{M}$ *for all* $a, b \in \mathbb{R}$.

Next will be characterized measurable maps between a measurable space and a product of measurable spaces.

Theorem 19. *Let* $(\mathbf{X}, \mathcal{M})$ *be a measurable space and* $\{(\mathbf{X}_i, \mathcal{M}_i) : i \in I\}$ *be an indexed system of measurable spaces. Let* $\pi_j : \prod_{i\in I} \mathbf{X}_i \to \mathbf{X}_j$ *be the projection maps for each* $j \in I$. *Then a function* $f : \mathbf{X} \to \prod_{i\in I} \mathbf{X}_i$ *is* $\big(\mathcal{M}, \bigotimes_{i\in I} \mathcal{M}_i\big)$*-measurable if and only if* $f_j = \pi_j \circ f : \mathbf{X} \to \mathbf{X}_j$ *is* $(\mathcal{M}, \mathcal{M}_j)$*-measurable for all* $j \in I$.

Proof. Assume that $f : \mathbf{X} \to \prod_{i\in I} \mathbf{X}_i$ is $\big(\mathcal{M}, \bigotimes_{i\in I} \mathcal{M}_i\big)$-measurable. Let $j \in I$ and $E_j \in \mathcal{M}_j$. Then we have that

$$f_j^{-1}[E_j] = (\pi_j \circ f)^{-1}[E_j] = f^{-1}[\pi_j^{-1}[E_j]] \in \mathcal{M}$$

since f is $\big(\mathcal{M}, \bigotimes_{i\in I} \mathcal{M}_i\big)$-measurable and $\pi_j^{-1}[E_j] \in \bigotimes_{i\in I} \mathcal{M}_i$ by the definition of product σ-algebra. It follows that f_j is $(\mathcal{M}, \mathcal{M}_j)$-measurable. To prove the converse direction, assume that $f_j : \mathbf{X} \to \mathbf{X}_j$ is $(\mathcal{M}, \mathcal{M}_j)$-measurable for all $j \in I$. Recall that $\bigotimes_{i\in I} \mathcal{M}_i$ is generated by the collection

$$\{\pi_j^{-1}[E_j] : E_j \in \mathcal{M}_j,\ j \in I\}$$

Let $j \in I$ and $E_j \in M_j$. Then we have that

$$f^{-1}[\pi_j^{-1}[E_j]] = (\pi_j \circ f)^{-1}[E_j] = f_j^{-1}[E_j] \in \mathcal{M}$$

since f_j is $(\mathcal{M}, \mathcal{M}_j)$-measurable. It now follows from Proposition 5 that f is $\big(\mathcal{M}, \bigotimes_{i\in I} \mathcal{M}_i\big)$-measurable. □

As a corollary, we have that the products and sums of Borel measurable functions are Borel measurable.

Proposition 7. *Let* $(\mathbf{X}, \mathcal{M})$ *be a measurable space and* $f : \mathbf{X} \to \mathbb{R}$ *and* $g : \mathbf{X} \to \mathbb{R}$ *be measurable. Then* $f + g$ *and* $f \cdot g$ *are measurable.*

Proof. By Theorem 19, the map $H : \mathbf{X} \to \mathbb{R} \times \mathbb{R}$ defined by $x \mapsto (f(x), g(x))$ is $(\mathcal{M}, \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}))$-measurable. On the other hand, by Theorem 4, we have that $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2)$. Therefore f is $(\mathcal{M}, \mathcal{B}(\mathbb{R}^2))$-measurable.

Consider $Sum : \mathbb{R} \times \mathbb{R}$ defined by $(x, y) \mapsto x + y$ and $Prod : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined by $(x, y) \mapsto x \cdot y$. It is clear that Sum and $Prod$ are continuous functions with respect to the product topology on $\mathbb{R}^2$ and consequently, are $(\mathcal{B}(\mathbb{R}^2), \mathcal{B}(\mathbb{R}))$-measurable. It follows that $f + g = Sum \circ H$ and $f \cdot g = Prod \circ H$ are measurable. $\square$

3.2. **The Borel structure of $\overline{\mathbb{R}}$.** The *extended real numbers* is the set

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

where $-\infty$ and $+\infty$ are two additional "points" that are not in $\mathbb{R}$. The complete linear order structure on $\mathbb{R}$ extends to a complete linear order structure on $\overline{\mathbb{R}}$ in the obvious way: Declare $-\infty$ to be less than all real numbers and $+\infty$ to be greater than all real numbers. Endowed with this linear order structure, *every* subset of $\overline{\mathbb{R}}$ has a supremum and an infimum. The algebraic operations on $\mathbb{R}$ can also be (partially) extended to $\overline{\mathbb{R}}$ in a natural way with the exception of the case $\infty - \infty$ and with the convention that $0 \cdot \pm\infty = 0$. For more details regarding these, the reader is referred to [Fol99, Section 0.5]. In this subsection, we shall be more interested in the topological structure of $\overline{\mathbb{R}}$.

Consider the function $\rho : \overline{\mathbb{R}} \to [0, \infty)$ given by

$$\rho(x, y) = |\arctan(x) - \arctan(y)|$$

where the usual arctan function is extended to $\overline{\mathbb{R}}$ by $\arctan(+\infty) = \pi/2$ and $\arctan(-\infty) = -\pi/2$. Since arctan is one-to-one, we have that ρ is a metric. We next show that $\rho \restriction \mathbb{R}$ is a compatible metric with the Euclidean topology of $\mathbb{R}$.

Let $x \in \mathbb{R}$. Since $\arctan'(t) \leq 1$ for all $t \in \mathbb{R}$, by the Mean Value Theorem, we have that $\rho(x, y) \leq d(x, y) = |x - y|$ for every $x, y \in \mathbb{R}$. Consequently, for any $r \in \mathbb{R}^+$, we have that $B_d(x, r) \subseteq B_\rho(x, r)$. Moreover, since tan is continuous at the point $\arctan(x)$, for any $r \in \mathbb{R}^+$, there exists $\delta \in \mathbb{R}^+$ such that $B_\rho(x, \delta) \subseteq B_d(x, r)$.[2] Hence ρ and d are equivalent metrics and generate the same topology on $\mathbb{R}$.

Now consider the map $\varphi : \overline{\mathbb{R}} \to [\frac{-\pi}{2}, \frac{\pi}{2}]$ given by

$$\varphi(x) = \begin{cases} -\pi/2 & \text{if } x = -\infty \\ \arctan(x) & \text{if } -\infty < x < \infty \\ \pi/2 & \text{if } x = +\infty \end{cases}$$

It is readily verified φ is a homeomorphism where $\overline{\mathbb{R}}$ is endowed with the topology induced by the metric ρ and $[\frac{-\pi}{2}, \frac{\pi}{2}]$ is endowed with its subspace topology coming from the Euclidean topology of $\mathbb{R}$, with which it is a compact space. Therefore, $\overline{\mathbb{R}}$ is a compact topological space and indeed, is a two-point compactification of $\mathbb{R}$.

Let us now consider the Borel σ-algebra of $\overline{\mathbb{R}}$. We claim that

$$\begin{aligned} \mathcal{B}(\overline{\mathbb{R}}) &= \{\ A \cup B :\ A \in \mathcal{B}(\mathbb{R}),\ B \subseteq \{+\infty, -\infty\}\ \} \\ &= \{\ S \subseteq \overline{\mathbb{R}} :\ S \cap \mathbb{R} \in \mathcal{B}(\mathbb{R})\ \} \end{aligned}$$

[2]This is the δ-value corresponding to $\epsilon = r$ in the ϵ-δ definition of continuity of tan at the point $\arctan(x)$.

Since every open ball in $\mathbb{R}$ is also an open set in $\overline{\mathbb{R}}$, we have that $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{B}(\overline{\mathbb{R}})$. Moreover, any subset of $\{+\infty, -\infty\}$ is a closed set in $\overline{\mathbb{R}}$ and hence is in $\mathcal{B}(\overline{\mathbb{R}})$. As $\mathcal{B}(\overline{\mathbb{R}})$ is closed under unions, the right-hand side is a subset of $\mathcal{B}(\overline{\mathbb{R}})$. To see the converse inclusion, observe that the collection $\{B_\rho(x,r) : x \in \overline{\mathbb{R}},\ r \in \mathbb{R}^+\}$ is a set of generators for $\mathcal{B}(\overline{\mathbb{R}})$ and every set in this collection is either an open subset of $\mathbb{R}$ or is of the form $O \cup \{\pm\infty\}$ for some open subset O of $\mathbb{R}$. It is easily seen that the right-hand side is a σ-algebra and consequently, contains $\mathcal{B}(\overline{\mathbb{R}})$ as a subset.

The reader may verify that $\mathcal{B}(\overline{\mathbb{R}})$ is also generated by the collections

$$\{[-\infty, a) : a \in \overline{\mathbb{R}}\} \quad \text{and} \quad \{(a, \infty] : a \in \overline{\mathbb{R}}\}$$

separately. Unless stated otherwise, $\overline{\mathbb{R}}$ is always endowed with this Borel structure.

3.3. **Measurable functions as limits of simple functions.** Next shall be shown that pointwise limits of $\overline{\mathbb{R}}$-valued measurable functions are measurable.

Theorem 20. *Let $(\mathbf{X}, \mathcal{M})$ be a measurable space and let $f_n : \mathbf{X} \to \overline{\mathbb{R}}$ be measurable for every $n \in \mathbb{N}$. Then the functions $F, G, H, K : \mathbf{X} \to \overline{\mathbb{R}}$ defined by*

- $F(x) = \sup_{n\in\mathbb{N}} f_n(x)$,
- $G(x) = \inf_{n\in\mathbb{N}} f_n(x)$,
- $H(x) = \limsup_{n\to\infty} f_n(x)$ *and*
- $K(x) = \liminf_{n\to\infty} f_n(x)$

for all $x \in \mathbf{X}$, are measurable.

Proof. Recall that $\mathcal{B}(\overline{\mathbb{R}})$ is generated by the sets of the form $(a, \infty]$ with $a \in \mathbb{R}$. Thus, by Proposition 5, it suffices to check that the inverse images of sets of this form are measurable. Let $a \in \mathbb{R}$. It easily follows from the definition of supremum that

$$F^{-1}[(a,\infty]] = \bigcup_{n\in\mathbb{N}} f_n^{-1}[(a,\infty]]$$

Since f_n's are measurable, we have that $f_n^{-1}[(a,\infty]] \in \mathcal{M}$ for every $n \in \mathbb{N}$ and hence, $F^{-1}[(a,\infty]] \in \mathcal{M}$ as $\mathcal{M}$ is a σ-algebra. Therefore, F is measurable. Using the sets of the form $[-\infty, a)$ which also generate $\overline{\mathbb{R}}$, one can also show with a similar trick that G is measurable.

Having shown that the pointwise supremum and infimum of a countable set of measurable functions are measurable, we have that $H_n(x) = \sup_{k\geq n} f_k(x)$ and $K_n(x) = \inf_{k\geq n} f_k(x)$ are measurable for all $n \in \mathbb{N}$. Consequently, H and K are measurable since $H(x) = \inf_{n\in\mathbb{N}} H_n(x)$ and $K(x) = \sup_{n\in\mathbb{N}} K_n(x)$. □

Theorem 20 has some important corollaries. First, given a sequence of $\overline{\mathbb{R}}$-valued measurable functions $(f_n)_{n\in\mathbb{N}}$ on a measurable space $(\mathbf{X}, \mathcal{M})$, if the pointwise limit $f(x) = \lim_{n\to\infty} f_n(x)$ exists for every $x \in \mathbf{X}$, then $f : \mathbf{X} \to \overline{\mathbb{R}}$ is measurable. Second, if $f : \mathbf{X} \to \overline{\mathbb{R}}$ and $g : \mathbf{X} \to \overline{\mathbb{R}}$ are measurable functions, then the maps $M(x) = \max\{f(x), g(x)\}$ and $m(x) = \min\{f(x), g(x)\}$ are measurable.

Next will be introduced a notion of utmost importance, namely, the notion of a simple function. Let $(\mathbf{X}, \mathcal{M})$ be a measurable space. A function $f : \mathbf{X} \to \mathbb{R}$ is said to be a *simple function* if it is measurable and has finite range.

Let $f : \mathbf{X} \to \mathbb{R}$ be a simple function.[3] By assumption, its range is finite, say, $ran(f) = \{r_1, r_2, \ldots, r_n\}$ with $r_i \neq r_j$ for distinct i and j. Since f is measurable, the sets $A_i = f^{-1}[\{r_i\}]$ are in $\mathcal{M}$ for every $1 \leq i \leq n$. Moreover, we have that

$$f(x) = \sum_{i=1}^{n} r_i \cdot \chi_{A_i}(x)$$

for every $x \in \mathbf{X}$. The expression above is called the *standard representation* of the simple function f. As we have just seen, any simple function is a finite linear combination of characteristic functions of measurable sets and conversely, any such function is simple.[4] For this reason, simple functions are often defined as finite linear combinations of characteristic functions of measurable sets in many books.

The reason that simple functions are central to the development of our theory of integration is that, as will be proven in the next theorem, every positive measurable function is the pointwise limit of simple functions. Consequently, in order to prove that measurable functions possess a certain property, one usually proves that simple functions have that property and it is preserved under taking pointwise limits.

Theorem 21. *Let $(\mathbf{X}, \mathcal{M})$ be a measurable space and let $f : \mathbf{X} \to [0, \infty]$ be measurable. Then there exists a sequence $(\phi_n)_{n\in\mathbb{N}}$ of simple functions such that $0 \leq \phi_n \leq \phi_{n+1} \leq f$ for all $n \in \mathbb{N}$ and $\lim_{n\to\infty} \phi_n(x) = f(x)$ for all $x \in \mathbf{X}$.*

Proof. Let $n \in \mathbb{N}$. For each integer $0 \leq k \leq 2^{2n} - 1$, set

$$A_k^n = f^{-1}\left[\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right)\right]$$

and $B^n = f^{-1}[[2^n, \infty]]$. As f is measurable, A_k^n's and B^n are in $\mathcal{M}$. Set $\phi_n : \mathbf{X} \to \overline{\mathbb{R}}$ to be the simple function given by

$$\phi_n(x) = \sum_{k=0}^{2^{2n}-1} \frac{k}{2^n} \cdot \chi_{A_k^n}(x) + 2^n \cdot \chi_{B^n}(x)$$

It is fairly straightforward to check that $\phi_n \leq f$. We next show that $\phi_n \leq \phi_{n+1}$ for all $n \in \mathbb{N}$. Let $n \in \mathbb{N}$ and $x \in \mathbf{X}$. Assume that $f(x) \in \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right)$ for some $0 \leq k \leq 2^{2n} - 1$. Then $f(x) \in \left[\frac{2k}{2^{n+1}}, \frac{2k+2}{2^{n+1}}\right)$ and hence

$$\phi_n(x) = \frac{k}{2^n} \leq \frac{2k}{2^{n+1}} \leq \phi_{n+1}(x)$$

[3]We would like to emphasize that, following the general convention, simple functions have codomain $\mathbb{R}$ and are not allowed to take the values $\pm\infty$.

[4]Note that a simple function may be represented in multiple ways as a finite linear combination of characteristic functions of measurable sets, however, the standard representation is the unique one where the coefficients are distinct and the measurable sets corresponding to the characteristic functions form a partition of the space.

Assume that $f(x) \in [2^n, \infty]$. Then $f(x) \in [2^{n+1}, \infty]$ or $f(x) \in \left[\frac{k}{2^{n+1}}, \frac{k+1}{2^{n+1}}\right)$ for some $2^{2n+1} \leq k \leq 2^{2n+2} - 1$. In both cases, we have that

$$\phi_n(x) = 2^n = \frac{2^{2n+1}}{2^{n+1}} \leq \phi_{n+1}(x)$$

It follows that $\phi_n(x) \leq \phi_{n+1}(x)$ for all $x \in \mathbf{X}$. Finally, we will prove that $\lim_{n\to\infty} \phi_n(x) = f(x)$ for all $x \in \mathbf{X}$.

Let $x \in \mathbf{X}$. We split into two cases. Assume that $f(x) \neq \infty$. Then, for sufficiently large values of n, we have that $x \in A_k^n$ for some $0 \leq k \leq 2^{2n} - 1$. Moreover, if $x \in A_k^n$, then we have that

$$|f(x) - \phi_n(x)| \leq \frac{1}{2^n}$$

It follows that $\lim_{n\to\infty} \phi_n(x) = f(x)$. Assume that $f(x) = \infty$. Then, by definition, $\lim_{n\to\infty} \phi_n(x) = \lim_{n\to\infty} 2^n = \infty = f(x)$. ☐

Analyzing the last estimation in the proof of Theorem 21, we indeed see that the sequence $(\phi_n)_{n\in\mathbb{N}}$ that we constructed converges to f *uniformly* on every set where f is bounded.

While carrying out proofs, we shall often need to modify measurable functions on μ-null sets. The next proposition shows that this trick does not disturb the measurability of the function if the measure space is complete. More precisely, we have the following.

Proposition 8. *Let $(\mathbf{X}, \mathcal{M}, \mu)$ be a complete measure space and $(\mathbf{Y}, \mathcal{N})$ be a measurable space. Let $f, g : \mathbf{X} \to \mathbf{Y}$ be functions. If f is measurable and $f(x) = g(x)$ holds μ-almost everywhere, then g is measurable.*

Proof. Assume that f is measurable and $N = \{x \in \mathbf{X} : f(x) \neq g(x)\}$ is μ-null. Let $E \in \mathcal{N}$. Since f is measurable, $f^{-1}[E] \in \mathcal{M}$ and so

$$g^{-1}[E] \cap N^c = f^{-1}[E] \cap N^c \in \mathcal{M}$$

As $(\mathbf{X}, \mathcal{M}\, \mu)$ is complete and N is μ-null, we have $g^{-1}[E] \cap N \in \mathcal{M}$. It follows that

$$g^{-1}[E] = (g^{-1}[E] \cap N) \cup (g^{-1}[E] \cap N^c) \in \mathcal{M}$$

Therefore, g is measurable. ☐

Corollary 22. *Let $(\mathbf{X}, \mathcal{M}\, \mu)$ be a complete measure space and $f_n : \mathbf{X} \to \overline{\mathbb{R}}$ be measurable for every $n \in \mathbb{N}$ and $f : \mathbf{X} \to \overline{\mathbb{R}}$ be a function. If $\lim_{n\to\infty} f_n(x) = f(x)$ μ-almost everywhere, then f is measurable.*

Proof. Assume that $\lim_{n\to\infty} f_n(x) = f(x)$ μ-almost everywhere. Set

$$N = \{x \in \mathbf{X} : \lim_{n\to\infty} f_n(x) \neq f(x)\}$$

By assumption, $\mu(N) = 0$. For each $n \in \mathbb{N}$, define the function $g_n = f_n \cdot \chi_{\mathbf{X}-N}$. Then g_n's are measurable by Proposition 7 and $\lim_{n\to\infty} g_n(x) = f(x) \cdot \chi_{\mathbf{X}-N}(x)$

for *all* $x \in \mathbf{X}$. Therefore, by Theorem 20, $f \cdot \chi_{\mathbf{X}-N}$ is measurable. On the other hand, $f(x)\chi_{\mathbf{X}-N}(x) = f(x)$ holds μ-almost everywhere and hence f is measurable by Proposition 8. □

3.4. **Integrating measurable functions.** In this subsection, we shall define the integral of a measurable function with respect to a measure. For the remaining of this subsection, let $(\mathbf{X}, \mathcal{M}, \mu)$ be a fixed measure space and let

$$L^+(\mathbf{X}, \mathcal{M}, \mu)^5 = \{f : \mathbf{X} \to [0, \infty] \ : \ f \text{ is measurable}^6\}$$

Our first goal is to define the integral of functions in $L^+(\mathbf{X}, \mathcal{M}, \mu)$. We know by Theorem 21 that the elements of $L^+(\mathbf{X}, \mathcal{M}, \mu)$ are pointwise limits of positive simple functions. For this reason, the first step will be to define the integral of simple functions and then try to extend this definition to functions in $L^+(\mathbf{X}, \mathcal{M}, \mu)$.

Let $\phi \in L^+(\mathbf{X}, \mathcal{M}, \mu)$ be a simple function with the standard representation

$$\phi(x) = \sum_{i=1}^{n} r_i \cdot \chi_{A_i}(x)$$

We define *the integral of ϕ with respect to μ* to be

$$\int_{\mathbf{X}} \phi \, d\mu = \sum_{i=1}^{n} r_i \cdot \mu(A_i)$$

with the convention that $0 \cdot \infty = 0$. Although we have defined $\int_{\mathbf{X}} \phi \, d\mu$ using the standard representation of ϕ, one may wish to make the same definition for an *arbitrary* representation of ϕ as a finite linear combination of characteristic functions of measurable sets, which may possibly include zero efficients or characteristic functions of non-disjoint measurable sets or empty sets. It is slightly cumbersome to check that the integral $\int_{\mathbf{X}} \phi \, d\mu$ is indeed independent of the representation of ϕ. While we are planning to use this fact, we shall not prove it here. For a proof of this fact, the reader may check [SS05, §2 Proposition 1.1]. We next define the integral of ϕ over an arbitrary measurable set. For each $A \in \mathcal{M}$, we define *the integral of ϕ with respect to μ over A* to be

$$\int_{A} \phi \, d\mu = \int_{\mathbf{X}} \phi \cdot \chi_A \, d\mu$$

Note that this definition makes sense since $\phi \cdot \chi_A$ is simple whenever ϕ is simple. We next show that the integral we defined has its expected linearity and monotonicity properties.

Lemma 3. *Let $\phi, \psi \in L^+(\mathbf{X}, \mathcal{M}, \mu)$ be simple functions. Then*

---

[5] We would like to remark that this set does **not** depend on the measure μ but depends only on the measurable space structure of $(\mathbf{X}, \mathcal{M}, \mu)$. For this reason, we should really have used the notation $L^+(\mathbf{X}, \mathcal{M})$ to denote it. However, since our purpose is the define the integral of functions in this set and the integral does depend on μ, we shall use this notation.

[6] Here $[0, \infty]$ is endowed with the Borel σ-algebra of its topology induced as a subspace of $\overline{\mathbb{R}}$.

a. $\int_{\mathbf{X}} c\phi\ d\mu = c\int_{\mathbf{X}} \phi\ d\mu$ *for every* $c \in \mathbb{R}^+$.
b. $\int_{\mathbf{X}} \phi + \psi\ d\mu = \int_{\mathbf{X}} \phi\ d\mu + \int_{\mathbf{X}} \psi\ d\mu$.
c. *If* $\phi \leq \psi$, *then* $\int_{\mathbf{X}} \phi\ d\mu \leq \int_{\mathbf{X}} \psi\ d\mu$.

Proof. Let $\phi = \sum_{i=1}^{n} r_i \cdot \chi_{A_i}$ and $\psi = \sum_{j=1}^{m} s_j \cdot \chi_{B_j}$ be the standard representations of ϕ and ψ. Recall that this means that the collections $\{A_i\}_{i=1}^{n}$ and $\{B_j\}_{j=1}^{m}$ are partitions of $\mathbf{X}$, the coefficients r_i's are distinct and the coefficients s_j's are distinct. Then we have $= \sum_{i=1}^{n}(cr_i) \cdot \chi_{A_i}$ is the standard representation of $c\phi$, from which part a follows.

To prove part b, set $C_{ij} = A_i \cap B_j$. It is immediately seen that C_{ij}'s are disjoint and moreover, $A_i = \bigsqcup_{j=1}^{m} C_{ij}$ and $B_j = \bigsqcup_{i=1}^{n} C_{ij}$. It follows that

$$\begin{aligned}
\int_{\mathbf{X}} \phi\ d\mu + \int_{\mathbf{X}} \psi\ d\mu &= \sum_{i=1}^{n} r_i \cdot \mu(A_i) + \sum_{j=1}^{m} s_j \cdot \mu(B_j) \\
&= \sum_{i=1}^{n} r_i \cdot \mu\left(\bigcup_{j=1}^{m} C_{ij}\right) + \sum_{j=1}^{m} s_j \cdot \mu\left(\bigcup_{i=1}^{n} C_{ij}\right) \\
&= \sum_{i=1}^{n} r_i \cdot \left(\sum_{j=1}^{m} \mu(C_{ij})\right) + \sum_{j=1}^{m} s_j \cdot \left(\sum_{i=1}^{n} \mu(C_{ij})\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m}(r_i + s_j)\mu(C_{ij}) = \int_{\mathbf{X}} \phi + \psi\ d\mu
\end{aligned}$$

We would like to emphasize that the last equality follows from that the integral of a simple function is independent of its representation and

$$\phi + \psi = \sum_{i=1}^{n}\sum_{j=1}^{m}(r_i + s_j)C_{ij}$$

is *some* representation of $\phi + \psi$.[7] Therefore, part b holds. To prove part c, observe that the finite linear combinations $\sum_{i=1}^{n}\sum_{j=1}^{m} r_i\chi_{C_{ij}}$ and $\sum_{j=1}^{m}\sum_{i=1}^{n} s_j\chi_{C_{ij}}$ are representations of ϕ and ψ respectively.[8] Consequently, if $\phi \leq \psi$, then $r_i \leq s_j$ whenever $C_{ij} \neq \emptyset$, in which case we have

$$\int_{\mathbf{X}} \phi\ d\mu = \sum_{i=1}^{n}\sum_{j=1}^{m} r_i \cdot \mu(C_{ij}) \leq \sum_{i=1}^{n}\sum_{j=1}^{m} s_j \cdot \mu(C_{ij}) = \int_{\mathbf{X}} \psi\ d\mu$$

This completes the proof. ∎

[7] The reason it may not be the standard representation is that the coefficients $r_i + s_j$ may fail to be distinct in which case we have to take the union of the corresponding measurable sets to regroup.

[8] This follows from the fact that if $E_1, \ldots, E_k$ are disjoint sets, then $\chi_{E_1 \sqcup \cdots \sqcup E_k} = \sum_{i=1}^{k} \chi_{E_k}$.

We are now ready to define the integral of functions in $L^+(\mathbf{X}, \mathcal{M}, \mu)$ over $\mathbf{X}$. Given $f \in L^+(\mathbf{X}, \mathcal{M}, \mu)$, we define

$$\int_{\mathbf{X}} f \, d\mu = \sup \left\{ \int_{\mathbf{X}} \phi \, d\mu : \ 0 \leq \phi \leq f, \ \phi \text{ is simple} \right\}$$

It follows from the definition of supremum that

$$\int_{\mathbf{X}} f \, d\mu \leq \int_{\mathbf{X}} g \, d\mu \text{ whenever } f \leq g$$

Moreover, Lemma 3.a implies that, for every $c \in \mathbb{R}^+$, we have

$$\int_{\mathbf{X}} cf \, d\mu = c \int_{\mathbf{X}} f \, d\mu$$

As before, for every measurable set $A \in \mathcal{M}$, we define

$$\int_A f \, d\mu = \int_{\mathbf{X}} f \cdot \chi_A \, d\mu$$

which can also be checked to have the properties that we have mentioned above.

We will now prove the first of the three main convergence results that we shall learn in this course.

Theorem 23 (The Monotone Convergence Theorem)**.** *Let $f_n \in L^+(\mathbf{X}, \mathcal{M}, \mu)$ be such that $f_n \leq f_{n+1}$ for all $n \in \mathbb{N}$. Then we have that*

$$\int_{\mathbf{X}} \lim_{n\to\infty} f_n \, d\mu = \lim_{n\to\infty} \int_{\mathbf{X}} f_n \, d\mu$$

Proof. First, note that $f(x) = \sup_{n\in\mathbb{N}} f_n(x) = \lim_{n\to\infty} f_n(x)$ for every $x \in \mathbf{X}$ by the monotonicity assumption and hence, we have $\lim_{n\to\infty} f_n(x) \in L^+(\mathbf{X}, \mathcal{M}, \mu)$ by Theorem 20.

We clearly have $f_k \leq f$ and so, $\int f_k \leq \int f$ for every $k \in \mathbb{N}$. It follows that

$$\lim_{n\to\infty} \int_{\mathbf{X}} f_n \, d\mu = \sup_{n\in\mathbb{N}} \int_{\mathbf{X}} f_n \, d\mu \leq \int_{\mathbf{X}} f \, d\mu$$

For the reverse inequality, let $0 \leq \phi \leq f$ be a simple function and let $0 < \alpha < 1$. set

$$E_n = \{x \in \mathbf{X} : f_n(x) \geq \alpha\phi(x)\}$$

for each $n \in \mathbb{N}$. Since $f(x) = \sup_{n\in\mathbb{N}} f_n(x)$ for all $x \in \mathbf{X}$ and $\alpha\phi(x) < f(x)$ unless $f(x) = 0$, in which case $x \in E_0$, we have that $\mathbf{X} = \bigcup_{n\in\mathbb{N}} E_n$ and moreover, $E_n \subseteq E_{n+1}$ for all $n \in \mathbb{N}$ by the monotonicity assumption. In order to proceed, we will need the following lemma.

Lemma 4. *If $\psi : \mathbf{X} \to [0, \infty]$ is simple, then $E \mapsto \int_E \psi \, d\mu$ is a measure on $(\mathbf{X}, \mathcal{M})$.*

Proof. Clearly $\int_\emptyset \psi \, d\mu = 0$. Let $A_1, A_2, \dots \in \mathcal{M}$ be disjoint. Let $\psi = \sum_{k=1}^n r_k \cdot \chi_{E_k}$ be the standard representation of ψ. Then we have that

$$\begin{aligned}
\int_{\bigsqcup_{i=1}^\infty A_i} \psi \, d\mu &= \int_{\mathbf{X}} \chi_{\bigsqcup_{i=1}^\infty A_i} \cdot \sum_{k=1}^n r_k \cdot \chi_{E_k} \, d\mu \\
&= \int_{\mathbf{X}} \sum_{k=1}^n r_k \cdot \chi_{E_k \cap \bigsqcup_{i=1}^\infty A_i} \, d\mu \\
&= \sum_{k=1}^n r_k \cdot \mu\left(E_k \cap \bigsqcup_{i=1}^\infty A_i\right) \\
&= \sum_{k=1}^n r_k \cdot \sum_{i=1}^\infty \mu\left(E_k \cap A_i\right) \\
&= \sum_{k=1}^n \sum_{i=1}^\infty r_k \cdot \mu\left(E_k \cap A_i\right) = \sum_{i=1}^\infty \sum_{k=1}^n r_k \cdot \mu\left(E_k \cap A_i\right) \\
&= \sum_{i=1}^\infty \int_{\mathbf{X}} r_k \cdot \chi_{E_k \cap A_i} \, d\mu = \sum_{i=1}^\infty \int_{A_i} r_k \cdot \chi_{E_k} = \sum_{i=1}^\infty \int_{A_i} \psi \, d\mu
\end{aligned}$$

It follows that the map $E \mapsto \int_E \psi \, d\mu$ is a measure. □

We now return to the proof of the Monotone Convergence Theorem. It is easily seen that

$$\int_{\mathbf{X}} f_n \, d\mu \geq \int_{E_n} f_n \, d\mu \geq \int_{E_n} \alpha\phi \, d\mu$$

On the other hand, since the map $E \mapsto \int_E \alpha\phi \, d\mu$ is a measure and $E_0 \subseteq E_1 \subseteq \dots$, we have that

$$\alpha \int_{\mathbf{X}} \phi \, d\mu = \int_{\mathbf{X}} \alpha\phi \, d\mu = \int_{\bigcup_{n\in\mathbb{N}} E_n} \alpha\phi \, d\mu = \lim_{n\to\infty} \int_{E_n} \alpha\phi \, d\mu \leq \lim_{n\to\infty} \int_{\mathbf{X}} f_n \, d\mu$$

The above inequality is true for all $0 < \alpha < 1$ and hence it is true for $\alpha = 1$. This implies that

$$\int_{\mathbf{X}} \phi \, d\mu \leq \lim_{n\to\infty} \int_{\mathbf{X}} f_n \, d\mu$$

As $0 \leq \phi \leq f$ was arbitrary, we have that

$$\int_{\mathbf{X}} f \, d\mu = \sup\left\{\int_{\mathbf{X}} \phi \, d\mu : \ 0 \leq \phi \leq f, \ \phi \text{ is simple}\right\} \leq \lim_{n\to\infty} \int_{\mathbf{X}} f_n \, d\mu$$

which finishes the proof of the Monotone Convergence Theorem. □

We would like to remark that the monotonicity assumption in this theorem cannot be dropped. For example, if $f_n = \chi_{[n,n+1)}$ for every $n \in \mathbb{N}$, then we have that $\lim_{n\to\infty} f_n(x) = 0$ for every $x \in \mathbb{R}$ but

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n \, d\mathbf{m} = \lim_{n\to\infty} \mathbf{m}([n, n+1)) = 1 \neq 0 = \int_{\mathbb{R}} 0 \, d\mathbf{m} = \int_{\mathbb{R}} \lim_{n\to\infty} f_n \, d\mathbf{m}$$

However, as a consequence of Fatou's lemma, we shall prove a version of the Monotone Convergence Theorem later on, where the monotonicity assumption is replaced by an appropriate boundedness condition.

We will next see some consequence of the Monotone Convergence Theorem. First, in order to evaluate the integral $\int_{\mathbf{X}} f \ d\mu$ for a function $f \in L^+(\mathbf{X}, \mathcal{M}, \mu)$, one does not really need to take a supremum over a potentially uncountable set as the definition requires, but rather, has to only compute the limit $\lim_{n\to\infty} \int_{\mathbf{X}} \phi_n \ d\mu$ where $(\phi_n)_{n\in\mathbb{N}}$ is any increasing sequence of non-negative simple functions approaching to f pointwise.

Second, the integral we defined for non-negative measurable functions, as expected, is additive.

Proposition 9. *Let $f_n \in L^+(\mathbf{X}, \mathcal{M}, \mu)$ for every $n \in \mathbb{N}$. Then we have that $\int_{\mathbf{X}} \sum_{n=0}^{k} f_n \ d\mu = \sum_{n=0}^{k} \int_{\mathbf{X}} f_n \ d\mu$ for every $k \in \mathbb{N}$. Moreover, we have*

$$\int_{\mathbf{X}} \sum_{n=0}^{\infty} f_n \ d\mu = \sum_{n=0}^{\infty} \int_{\mathbf{X}} f_n \ d\mu$$

Proof. We prove the first claim by induction on $k \in \mathbb{N}$. The claim is trivial for $k = 0$. Let $k \in \mathbb{N}$ and assume that the claim holds for k. Let $(\phi_i)_{i\in\mathbb{N}}$ and $(\psi_i)_{i\in\mathbb{N}}$ be increasing sequences of simple functions in $L^+(\mathbf{X}, \mathcal{M}, \mu)$ approaching pointwise to $\sum_{n=0}^{k} f_n$ and f_{k+1} respectively.[9] Then, by various applications of the Monotone Convergence Theorem together with the induction assumption at the end, we have that

$$\begin{aligned}
\int_{\mathbf{X}} \sum_{n=0}^{k+1} f_n \ d\mu &= \int_{\mathbf{X}} \sum_{n=0}^{k} f_n + f_{k+1} \ d\mu \\
&= \int_{\mathbf{X}} \lim_{i\to\infty} (\phi_i + \psi_i) \ d\mu \\
&= \lim_{i\to\infty} \int_{\mathbf{X}} (\phi_i + \psi_i) \ d\mu \\
&= \lim_{i\to\infty} \left(\int_{\mathbf{X}} \phi_i \ d\mu + \int_{\mathbf{X}} \psi_i \ d\mu \right) = \lim_{i\to\infty} \int_{\mathbf{X}} \phi_i \ d\mu + \lim_{i\to\infty} \int_{\mathbf{X}} \psi_i \ d\mu \\
&= \int_{\mathbf{X}} \lim_{i\to\infty} \phi_i \ d\mu + \int_{\mathbf{X}} \lim_{i\to\infty} \psi_i \ d\mu \\
&= \int_{\mathbf{X}} \sum_{n=0}^{k} f_n \ d\mu + \int_{\mathbf{X}} f_{k+1} \ d\mu \\
&= \sum_{n=0}^{k} \int_{\mathbf{X}} f_n \ d\mu + \int_{\mathbf{X}} f_{k+1} \ d\mu = \sum_{n=0}^{k+1} \int_{\mathbf{X}} f_n \ d\mu
\end{aligned}$$

[9] Even though we have not proven it, the reader should show that the sum of functions in $L^+(\mathbf{X}, \mathcal{M}, \mu)$ is in $L^+(\mathbf{X}, \mathcal{M}, \mu)$. So such sequences of simple functions by Theorem 21.

Therefore, by induction, we have that the claim holds for all $k \in \mathbb{N}$. Clearly, the sequence $(g_k)_{k\in\mathbb{N}}$ of functions satisfy the hypotheses of the Monotone Convergence Theorem where $g_k = \sum_{n=0}^{k} f_n$ for each $k \in \mathbb{N}$. Thus, applying the Monotone Convergence Theorem once more and using the first part of this theorem, we get that

$$\int_{\mathbf{X}} \sum_{n=0}^{\infty} f_n \; d\mu = \int_{\mathbf{X}} \lim_{k\to\infty} g_k \; d\mu = \lim_{k\to\infty} \int_{\mathbf{X}} g_k \; d\mu = \lim_{k\to\infty} \sum_{n=0}^{k} \int_{\mathbf{X}} f_n \; d\mu = \sum_{n=0}^{\infty} \int_{\mathbf{X}} f_n \; d\mu$$

□

Another consequence of the Monotone Convergence Theorem is the following fact, which will be used later.

Proposition 10. *Let $f \in L^+(\mathbf{X}, \mathcal{M}, \mu)$. Then $\int_{\mathbf{X}} f \; d\mu = 0$ if and only if $f(x) = 0$ holds μ-almost everywhere.*

Proof. We will first show if $\phi \in L^+(\mathbf{X}, \mathcal{M}, \mu)$ is simple and $\phi(x) = 0$ holds μ-almost everywhere, then $\int_{\mathbf{X}} \phi \; d\mu = 0$. Let $\phi \in L^+(\mathbf{X}, \mathcal{M}, \mu)$ be a simple function with standard representation $\phi = \sum_{k=1}^{n} r_k \cdot \chi_{A_k}$ such that $\phi(x) = 0$ holds μ-almost everywhere. Then, a moment's thought reveals that, for every $1 \le k \le n$, we have that $\mu(A_k) = 0$ whenever $r_k \neq 0$. It follows that $\int_{\mathbf{X}} \phi \; d\mu = \sum_{k=1}^{n} r_k \cdot \mu(A_k) = 0$. We can now prove the proposition.

Assume that $f(x) = 0$ holds μ-almost everywhere. Let $0 \le \phi \le f$ be simple. Then $\phi(x) = 0$ holds μ-almost everywhere and hence $\int_{\mathbf{X}} \phi \; d\mu = 0$. Therefore, $\int_{\mathbf{X}} f \; d\mu = \sup\{\int_{\mathbf{X}} \phi \; d\mu : 0 \le \phi \le f, \; \phi \text{ is simple}\} = 0$.

Now assume that $f(x) = 0$ does not hold μ-almost everywhere. It follows that

$$\mu\left(\bigcup_{n=1}^{\infty} \left\{x \in \mathbf{X} : \; f(x) > \frac{1}{n}\right\}\right) = \mu\left(\{x \in \mathbf{X} : f(x) > 0\}\right) > 0$$

and hence $\mu\left(\left\{x \in \mathbf{X} : \; f(x) > \frac{1}{k}\right\}\right) > 0$ for some $k \in \mathbb{N}^+$ as a countable union of μ-null sets is μ-null. Consequently,

$$\int_{\mathbf{X}} f \; d\mu \ge \int_{\left\{x\in\mathbf{X} : \; f(x) > \frac{1}{k}\right\}} \frac{1}{k} = \mu\left(\left\{x \in \mathbf{X} : \; f(x) > \frac{1}{k}\right\}\right) \cdot \frac{1}{k} > 0$$

□

Next shall be proven the second convergence results that we shall learn in this class, namely, Fatou's lemma.

Theorem 24 (Fatou's lemma)**.** *Let $f_n \in L^+(\mathbf{X}, \mathcal{M}, \mu)$ for all $n \in \mathbb{N}$. Then*

$$\int_{\mathbf{X}} \liminf_{n\to\infty} f_n \; d\mu \le \liminf_{n\to\infty} \int_{\mathbf{X}} f_n \; d\mu$$

Proof. Let $k \in \mathbb{N}$. Then, for every integer $n \geq k$, we have $\inf_{n \geq k} f_n \leq f_k$ by definition of infimum. It follows that

$$\int_{\mathbf{X}} \inf_{n \geq k} f_n \ d\mu \leq \int_{\mathbf{X}} f_k \ d\mu$$

Since this is true for each $k \in \mathbb{N}$, we have that

$$\int_{\mathbf{X}} \inf_{n \geq k} f_n \ d\mu \leq \inf_{k \in \mathbb{N}} \int_{\mathbf{X}} f_k \ d\mu$$

Applying the Monotone Convergence Theorem to the sequence $(\inf_{n \geq k} f_n)_{k \in \mathbb{N}}$ and using the previous inequality, we have that

$$\int_{\mathbf{X}} \liminf_{n \to \infty} f_n \ d\mu = \int_{\mathbf{X}} \lim_{k \to \infty} \inf_{n \geq k} f_n = \lim_{k \to \infty} \int_{\mathbf{X}} \inf_{n \geq k} f_n \leq \lim_{k \to \infty} \inf_{k \in \mathbb{N}} \int_{\mathbf{X}} f_k \ d\mu$$
$$\leq \liminf_{n \to \infty} \int_{\mathbf{X}} f_n \ d\mu$$

□

We will now prove an important corollary of Fatou's lemma, which may be considered as a "cousin" of the Monotone Convergence Theorem. We would like to note that, within the proof of this fact, we shall carry out an important trick that will also be used in our arguments later, namely, modifying functions on null sets to be able to apply our tools. The reader is expected to get used to this approach.

Theorem 25. *Let $f_n, f \in L^+(\mathbf{X}, \mathcal{M}, \mu)$ be such that $f_n(x) \longrightarrow f(x)$ holds μ-almost everywhere and $f_n(x) \leq f(x)$ holds μ-almost everywhere for each $n \in \mathbb{N}$. Then*

$$\lim_{n \to \infty} \int_{\mathbf{X}} f_n \ d\mu = \int_{\mathbf{X}} f \ d\mu$$

Proof. Let $N = \{x \in \mathbf{X} : f_n(x) \nrightarrow f(x), \text{ or, } f_n(x) \nleq f(x) \text{ for some } n \in \mathbb{N}\}$. By the hypothesis and the fact that a countable union of μ-null sets is μ-null, we see that $\mu(N) = 0$. Set $g_n = f_n \cdot \chi_{N^c}$ for each $n \in \mathbb{N}$ and set $g = f \cdot \chi_{N^c}$. It is easily checked that we have $g_n(x) \to g(x)$ and $g_n(x) \leq g(x)$ for *every* $x \in \mathbf{X}$. Applying Fatou's lemma, we have that

$$\int_{\mathbf{X}} g \ d\mu = \int_{\mathbf{X}} \lim_{n \to \infty} g_n \ d\mu \leq \int_{\mathbf{X}} \liminf_{n \to \infty} g_n \ d\mu$$
$$\leq \liminf_{n \to \infty} \int_{\mathbf{X}} g_n \ d\mu \leq \limsup_{n \to \infty} \int_{\mathbf{X}} g_n \ d\mu \leq \int_{\mathbf{X}} g \ d\mu$$

Therefore $\lim_{n \to \infty} \int_{\mathbf{X}} g_n$ exists and $\lim_{n \to \infty} \int_{\mathbf{X}} g_n \ d\mu = \int_{\mathbf{X}} g \ d\mu$. On the other hand, $f_n - g_n = 0$ and $f - g = 0$ hold μ-almost everywhere and hence, by Proposition 10, we have that $\int f_n d\mu = \int g_n d\mu$ for every $n \in \mathbb{N}$ and $\int f d\mu = \int g d\mu$. Therefore

$$\lim_{n \to \infty} \int_{\mathbf{X}} f_n \ d\mu = \lim_{n \to \infty} \int_{\mathbf{X}} g_n \ d\mu = \int_{\mathbf{X}} g \ d\mu = \int_{\mathbf{X}} f \ d\mu$$

□

Having built a powerful theory of integration for non-negative measurable functions, we now extend this theory to all measurable functions as follows. Let

$$L(\mathbf{X}, \mathcal{M}, \mu) = \{f : \mathbf{X} \to \mathbb{R} \ : \ f \text{ is measurable}\}^{10}$$

Given a function $f \in L(\mathbf{X}, \mathcal{M}, \mu)$, we define its *positive part* f^+ and *negative part* f^- as follows.

$$f^+(x) = \max\{0, f(x)\} \text{ and } f^-(x) = \max\{0, -f(x)\}$$

for every $x \in \mathbf{X}$. It is easily checked that f^+ and f^- are in $L^+(\mathbf{X}, \mathcal{M}, \mu)$ and that

$$f = f^+ - f^- \text{ and } |f| = f^+ + f^-$$

We say that f *is integrable over* $\mathbf{X}$ if $\int_{\mathbf{X}} f^+ \, d\mu < \infty$ and $\int_{\mathbf{X}} f^- \, d\mu < \infty$. In this case, we define the *integral of* f *over* $\mathbf{X}$ to be

$$\int_{\mathbf{X}} f \, d\mu = \int_{\mathbf{X}} f^+ \, d\mu - \int_{\mathbf{X}} f^- \, d\mu$$

Assume that f is integrable over $\mathbf{X}$. Then $\int_{\mathbf{X}} f^+ \, d\mu$ and $\int_{\mathbf{X}} f^- \, d\mu$ are both finite. As the integral is additive for measurable functions in $L^+(\mathbf{X}, \mathcal{M}, \mu)$, we have that $\int |f| \, d\mu = \int f^+ + f^- \, d\mu = \int_{\mathbf{X}} f^+ \, d\mu + \int_{\mathbf{X}} f^- \, d\mu < \infty$. Conversely, assume that we have $\int |f| \, d\mu < \infty$. Then, $\int_{\mathbf{X}} f^+ \, d\mu$, $\int_{\mathbf{X}} f^- \, d\mu \le \int |f| \, d\mu < \infty$ as $f^+, f^- \le |f|$. In other words, we have shown that

$$f \text{ is integrable over } \mathbf{X} \text{ if and only if } \int_{\mathbf{X}} |f| \, d\mu < \infty$$

which may also be taken as a definition of integrability over $\mathbf{X}$. As before, for each measurable $E \in \mathcal{M}$, we say that f is *integrable over* E if $f \cdot \chi_E$ is integrable, in which case we define its integral to be

$$\int_E f \, d\mu = \int_{\mathbf{X}} f \cdot \chi_E \, d\mu$$

From now on, we set

$$\begin{aligned} L^1(\mathbf{X}, \mathcal{M}, \mu) &= \{f \in L(\mathbf{X}, \mathcal{M}, \mu) : \ f \text{ is integrable}\} \\ &= \left\{ f \in L(\mathbf{X}, \mathcal{M}, \mu) : \ \int_{\mathbf{X}} |f| \, d\mu < \infty \right\} \end{aligned}$$

The integral that we defined has its usual expected properties. A straightforward computation shows that it is linear, that is,

$$\int_E c \cdot f + g \, d\mu = c \int_E f \, d\mu + \int_E g \, d\mu$$

[10] We would like to remark that, as before, this set does **not** depend on the measure μ but depends only on the measurable space structure of $(\mathbf{X}, \mathcal{M})$. Still, we would keep use the notation $L(\mathbf{X}, \mathcal{M}, \mu)$ instead of $L(\mathbf{X}, \mathcal{M})$. Also, the codomain $\mathbb{R}$ is endowed with its Borel structure $\mathcal{B}(\mathbb{R})$.

for every $c \in \mathbb{R}$, every $E \in \mathcal{M}$ and every $f, g \in L^1(\mathbf{X}, \mathcal{M}, \mu)$. Moreover, the triangle inequality for integrals is satisfied as

$$\left|\int_{\mathbf{X}} f \; d\mu\right| = \left|\int_{\mathbf{X}} f^+ - f^- \; d\mu\right| \leq \int_{\mathbf{X}} f^+ \; d\mu + \int_{\mathbf{X}} f^- d\mu = \int_{\mathbf{X}} f^+ + f^- d\mu = \int_{\mathbf{X}} |f| \; d\mu$$

Recall that two non-negative measurable functions that are almost everywhere have the same integral. The same result holds for integrable functions.

Lemma 5. *Let $f, g \in L^1(\mathbf{X}, \mathcal{M}, \mu)$. Then the following are equivalent.*

a. $\int_E f \; d\mu = \int_E g \; d\mu$ *for every* $E \in \mathcal{M}$.
b. $\int_{\mathbf{X}} |f - g| \; d\mu = 0$.
c. $f(x) = g(x)$ *holds μ-almost everywhere.*

Proof. Since $|f - g| \in L^+(\mathbf{X}, \mathcal{M}, \mu)$, Proposition 10 implies that $b \Leftrightarrow c$. We now show that $b \Rightarrow a$. Assume that $\int_{\mathbf{X}} |f - g| \; d\mu = 0$. Let $E \in \mathcal{M}$. Then we have that

$$\left|\int_E f \; d\mu - \int_E g \; d\mu\right| = \left|\int_E f - g \; d\mu\right| \leq \left|\int_{\mathbf{X}} f - g \; d\mu\right| \leq \int_{\mathbf{X}} |f - g| \; d\mu = 0$$

Therefore $\int_E f \; d\mu = \int_E g \; d\mu$. We next shot that $a \Rightarrow c$, which would complete the proof. Assume $\int_E f \; d\mu = \int_E g \; d\mu$ for every $E \in \mathcal{M}$ and suppose towards a contradiction that $f(x) = g(x)$ does not μ-almost everywhere. Then

$$\mu\left(\{x \in \mathbf{X} : \; f(x) \neq g(x)\}\right) > 0$$
$$\mu\left(\{x \in \mathbf{X} : \; f^+(x) - f^-(x) \neq g^+(x) - g^-(x)\}\right) > 0$$
$$\mu\left(\{x \in \mathbf{X} : \; f^+(x) - g^+(x) \neq f^-(x) - g^-(x)\}\right) > 0$$
$$\mu\left(\{x \in \mathbf{X} : \; (f - g)^+(x) \neq (f - g)^-(x)\}\right) > 0$$
$$\mu\left(\{x \in \mathbf{X} : \; (f - g)^+(x) > 0\}\right) + \mu\left(\{x \in \mathbf{X} : \; (f - g)^-(x) > 0\}\right) > 0$$

Without loss of generality, we may assume that $\mu\left(\{x \in \mathbf{X} : \; (f - g)^+(x) > 0\}\right) > 0$. Then, since strictly positive functions have strictly positive integrals on positive measure sets, we have that

$$\int_{\{x \in \mathbf{X} : \; (f-g)^+(x) > 0\}} f - g \; d\mu = \int_{\{x \in \mathbf{X} : \; (f-g)^+(x) > 0\}} (f - g)^+ \; d\mu > 0$$

which contradicts our assumption. □

Lemma 5 shows that μ-almost everywhere equal functions have the same integral. Thus, for the purposes of integration, it suffices to consider functions up to μ-almost everywhere equivalence.

Let $\sim$ be the equivalence relation on $L^1(\mathbf{X}, \mathcal{M}, \mu)$ given by

$$f \sim g \iff f(x) = g(x) \text{ holds } \mu\text{-almost everywhere}$$

Consider the quotient space

$$\mathbf{L}^1(\mathbf{X}, \mathcal{M}, \mu) = L^1(\mathbf{X}, \mathcal{M}, \mu) / \sim$$

together with the function $\rho : \mathbf{L}^1(\mathbf{X}, \mathcal{M}, \mu) \to \mathbb{R}$ given by

$$\rho([f],[g]) = \int_{\mathbf{X}} |f-g| \; d\mu$$

Then, it follows from Lemma 5 that the pair $(\mathbf{L}^1(\mathbf{X}, \mathcal{M}, \mu), \rho)$ is a metric space. We will later investigate the relationship between convergence of functions in this space and convergence in other senses. $\mathbf{L}^1(\mathbf{X}, \mathcal{M}, \mu)$ is indeed a Banach space together with the norm $\int_{\mathbf{X}} |\cdot| \; d\mu$.

It is now time for us to prove the last of the three convergence theorems, namely, Lebesgue's dominated convergence theorem.

Theorem 26 (Lebesgue's dominated convergence theorem)**.** *Let $f_n \in L(\mathbf{X}, \mathcal{M}, \mu)$ for all $n \in \mathbb{N}$ and let $f \in L(\mathbf{X}, \mathcal{M}, \mu)$ be such that*

- *$f_n(x) \longrightarrow f(x)$ holds μ-almost everywhere.*
- *There exists a non-negative integrable function $g \in L^1(\mathbf{X}, \mathcal{M}, \mu)$ such that $|f_n| \leq g$ holds μ-almost everywhere for every $n \in \mathbb{N}$.*

Then f_n's and f are integrable and we have that

$$\lim_{n\to\infty} \int_{\mathbf{X}} f_n \; d\mu = \int_{\mathbf{X}} f \; d\mu$$

Proof. Set

$$K = \mathbf{X} - \left(\{x \in \mathbf{X} : f_n(x) \nrightarrow f(x)\} \cup \bigcup_{n\in\mathbb{N}} \{x \in \mathbf{X} : |f_n(x)| > g(x)\} \right)$$

By assumption, each of the sets on the right hand side are μ-null and hence K^c is μ-null. On the other hand, for every $x \in K$, we have that $|f_n(x)| \leq g(x)$ for all $n \in \mathbb{N}$ and $\lim_{n\to\infty} f_n(x) = f(x)$, which together imply that $|f(x)| \leq g(x)$. It follows that

$$\int_{\mathbf{X}} |f| \; d\mu = \int_K |f| \; d\mu \leq \int_K g \; d\mu \leq \int_{\mathbf{X}} |g| \; d\mu < \infty$$

Similarly, one has $\int_{\mathbf{X}} |f_n| \; d\mu \leq \int_{\mathbf{X}} |g| \; d\mu < \infty$ for all $n \in \mathbb{N}$. Therefore, each f_n and f are integrable. Note that $g - f_n \geq 0$ and $g + f_n \geq 0$ for every $x \in K$. Applying Fatou's lemma to the sequence $(g + f_n)_{n\in\mathbb{N}}$ over K, we have that

$$\begin{aligned}
\int_K g \; d\mu + \int_K f \; d\mu = \int_K (g+f) \; d\mu &= \int_K \liminf_{n\to\infty} (g+f_n) \; d\mu \\
&\leq \liminf_{n\to\infty} \left(\int_K g \; + f_n \; d\mu \right) \\
&\leq \liminf_{n\to\infty} \left(\int_K g \; d\mu + \int_K f_n \; d\mu \right) \\
&\leq \int_K g \; d\mu + \liminf_{n\to\infty} \int_K f_n \; d\mu
\end{aligned}$$

Similarly, applying Fatou's lemma to the sequence $(g-f_n)_{n\in\mathbb{N}}$ over K, we get

$$\begin{aligned}\int_K g\ d\mu - \int_K f\ d\mu = \int_K (g-f)\ d\mu = \int_K \liminf_{n\to\infty}(g-f_n)\ d\mu \\ &\leq \liminf_{n\to\infty}\left(\int_K g\ - f_n\ d\mu\right) \\ &\leq \liminf_{n\to\infty}\left(\int_K g\ d\mu - \int_K f_n\ d\mu\right) \\ &\leq \int_K g\ d\mu + \liminf_{n\to\infty}\left(-\int_K f_n\ d\mu\right) \\ &\leq \int_K g\ d\mu - \limsup_{n\to\infty}\int_K f_n\ d\mu\end{aligned}$$

Consequently,

$$\int_K f\ d\mu \leq \liminf_{n\to\infty}\int_K f_n\ d\mu \leq \limsup_{n\to\infty}\int_K f_n\ d\mu \leq \int_K f\ d\mu$$

from which it follows that $\lim_{n\to\infty}\int_K f_n\ d\mu$ exists and

$$\lim_{n\to\infty}\int_K f_n\ d\mu = \int_K f\ d\mu$$

On the other hand, as K^c is μ-null, it follows that

$$\lim_{n\to\infty}\int_{\mathbf{X}} f_n\ d\mu = \int_{\mathbf{X}} f\ d\mu < \infty$$

□

One may wish to weaken the hypotheses of the Dominated Convergence Theorem. We would like to note that some kind of domination hypotheses is necessary since $n\cdot\chi_{\left(0,\frac{1}{n}\right)} \to 0$ pointwise, however,

$$\lim_{n\to\infty}\int_{\mathbb{R}} n\cdot\chi_{\left(0,\frac{1}{n}\right)}\ d\mathbf{m} = \lim_{n\to\infty} 1 \neq 0 = \int_{\mathbb{R}} 0\ d\mathbf{m} = \int_{\mathbb{R}} \lim_{n\to\infty} n\cdot\chi_{\left(0,\frac{1}{n}\right)}\ d\mathbf{m}$$

As it was the case with other convergence theorems, the Dominated Convergence Theorem has some useful and important corollaries, some of which will be proven next. First, it allows us to interchange an integral and an infinite sum, provided that the infinite sum of the integrals of absolute values converges.

Theorem 27. *Let $f_n \in L(\mathbf{X},\mathcal{M},\mu)$ for all $n\in\mathbb{N}$ be such that*

$$\sum_{n=0}^{\infty}\int_{\mathbf{X}} |f_n|\ d\mu < \infty$$

Then we have

$$\int_{\mathbf{X}}\sum_{n=0}^{\infty} f_n\ d\mu = \sum_{n=0}^{\infty}\int_{\mathbf{X}} f_n\ d\mu$$

Proof. Clearly, each $|f_n|$ is integrable for the infinite sum $\sum_{n=0}^{\infty} \int_{\mathbf{X}} |f_n| \, d\mu$ would have diverged otherwise. Since $|f_n| \in L^+(\mathbf{X}, \mathcal{M}, \mu)$ for all $n \in \mathbb{N}$, by Proposition 9, we have that

$$\int_{\mathbf{X}} \sum_{n=0}^{\infty} |f_n| \, d\mu = \sum_{n=0}^{\infty} \int_{\mathbf{X}} |f_n| \, d\mu < \infty$$

Set $g(x) = \sum_{n=0}^{\infty} |f_n|(x)$ for all $x \in \mathbf{X}$. Then one can check that $g \in L^+(\mathbf{X}, \mathcal{M}, \mu)$. If $g(x) = \infty$ holds on a set of positive measure, then g would not have finite integral, and hence $g(x) < \infty$ holds μ-almost everywhere. Subsequently, $\sum_{n=0}^{\infty} f_n(x)$ converges to a finite value μ-almost everywhere, say, on a set $K \subseteq \mathbf{X}$ with $\mu(K^c) = 0$. Set $f = \sum_{n=0}^{\infty}(f_n \cdot \chi_K)$ and $g_k = \sum_{n=0}^{k}(f_n \cdot \chi_K)$ for each $k \in \mathbb{N}$. Then each g_k is measurable and consequently, f is measurable. Moreover, we have that $g_k \to f$ and $|g_k| \leq g \cdot \chi_K$ for each $k \in \mathbb{N}$. Thus, by the Dominated Convergence Theorem and the additivity of integral, we have that

$$\int_{\mathbf{X}} \sum_{n=0}^{\infty} f_n \, d\mu = \int_{\mathbf{X}} \lim_{k\to\infty} g_k \, d\mu = \lim_{k\to\infty} \int_{\mathbf{X}} g_k \, d\mu = \lim_{k\to\infty} \sum_{n=0}^{k} \int_{\mathbf{X}} f_k \, d\mu = \sum_{n=0}^{\infty} \int_{\mathbf{X}} f_n \, d\mu$$

□

The next corollary is that the set of (equivalence classes) of simple functions are dense in $(\mathbf{L}^1(\mathbf{X}, \mathcal{M}, \mu), \rho)$.

Theorem 28. *For every integrable function $f \in L^1(\mathbf{X}, \mathcal{M}, \mu)$ and every $\epsilon \in \mathbb{R}^+$, there exists a simple function $\phi \in L^1(\mathbf{X}, \mathcal{M}, \mu)$ such that $\int_{\mathbf{X}} |f - \phi| \, d\mu < \epsilon$.*

Proof. Applying Theorem 21 to f^+ and f^-, we can find two increasing sequences $(\psi_n)_{n\in\mathbb{N}}$ and $(\eta_n)_{n\in\mathbb{N}}$ of non-negative simple functions such that $\psi_n \leq f^+$ and $\eta_n \leq f^-$ for all $n \in \mathbb{N}$. Set $\phi_n = \psi_n + \eta_n$. Then each $|f - \phi_n|$ is measurable, $|f - \phi_n| \to 0$ and $|f - \phi_n| \leq 2|f|$ for all $n \in \mathbb{N}$. It then follows from the Dominated Convergence Theorem that

$$0 = \int_X 0 \, d\mu = \int_{\mathbf{X}} \lim_{n\to\infty} |f - \phi_n| \, d\mu = \lim_{n\to\infty} \int_{\mathbf{X}} |f - \phi_n| \, d\mu$$

Therefore, there exists $k \in \mathbb{N}$ such that $\int_{\mathbf{X}} |f - \phi_k| \, d\mu < \epsilon$. □

It is often needed in Calculus to move a partial differentiation operator inside an integral. Such steps can be justified by the following theorem, which is also a corollary of the Dominated Convergence Theorem.

Theorem 29. *Let $f : \mathbf{X} \times (a, b) \to \mathbb{R}$ with $-\infty < a < b < \infty$ be such that*

- *$f_2(x, t)$ exists for all $t \in (a, b)$ and all $x \in \mathbf{X}$.*
- *There exists $g \in L^1(\mathbf{X}, \mathcal{M}, \mu)$ such that $|f_2(x, t)| \leq g(x)$ for all $t \in (a, b)$ and all $x \in \mathbf{X}$.*

Then

$$\frac{\partial}{\partial t} \int_{\mathbf{X}} f(x, t) \, d\mu = \int_{\mathbf{X}} \frac{\partial f(x, t)}{\partial t} \, d\mu$$

Proof. Let $\hat{t} \in (a,b)$ and let $(t_n)_{n\in\mathbb{N}}$ be a sequence of real numbers in (a,b) such that $\lim_{n\to\infty} t_n = \hat{t}$ and $t_n \neq \hat{t}$ for all $n \in \mathbb{N}$. Set

$$g_n(x) = \frac{f(x,t_n) - f(x,\hat{t})}{t_n - \hat{t}}$$

for all $x \in \mathbf{X}$. Then, by definition, for all $x \in \mathbf{X}$, we have

$$\lim_{n\to\infty} g_n(x) = f_2(x,\hat{t})$$

as this partial derivative exists at $(x,\hat{t})$. Note that g_n's are measurable, and being a pointwise limit of measurable functions, the map $f_2(\cdot,\hat{t}) : \mathbf{X} \to \mathbb{R}$ is measurable. Moreover, for every $x \in \mathbf{X}$, the map $\cdot \mapsto f(x,\cdot)$ is differentiable on any subinterval of (a,b) and hence, by the Mean Value Theorem, we obtain that

$$|g_n(x)| = \left|\frac{f(x,t_n) - f(x,\hat{t})}{t_n - \hat{t}}\right| = \sup_{u\in(t_n,\hat{t})} |f_2(x,u)| \leq g(x)$$

An application of the Dominated Convergence Theorem results in

$$\begin{aligned}
\lim_{n\to\infty} \frac{\int_{\mathbf{X}} f(x,t_n)\ d\mu - \int_{\mathbf{X}} f(x,\hat{t})\ d\mu}{t_n - \hat{t}} &= \lim_{n\to\infty} \int_{\mathbf{X}} \frac{f(x,t_n) - f(x,\hat{t})}{t_n - \hat{t}}\ d\mu \\
&= \lim_{n\to\infty} \int_{\mathbf{X}} g_n(x)\ d\mu \\
&= \int_{\mathbf{X}} \lim_{n\to\infty} g_n(x)\ d\mu \\
&= \int_{\mathbf{X}} \lim_{n\to\infty} \left.\frac{\partial f(x,t)}{\partial t}\right|_{t=\hat{t}} d\mu
\end{aligned}$$

Note that the limit on the left-hand side equals the right-hand side for every sequence $(t_n)_{n\in\mathbb{N}}$ with limit is $\hat{t}$ such that $t_n \neq \hat{t}$ for all $n \in \mathbb{N}$. It is now straightforward to verify that

$$\begin{aligned}
\left.\left(\frac{\partial}{\partial t}\int_{\mathbf{X}} f(x,t)\ d\mu\right)\right|_{t=\hat{t}} &= \lim_{h\to 0} \frac{\int_{\mathbf{X}} f(x,\hat{t}+h)\ d\mu - \int_{\mathbf{X}} f(x,\hat{t})\ d\mu}{h} \\
&= \int_{\mathbf{X}} \lim_{n\to\infty} \left.\frac{\partial f(x,t)}{\partial t}\right|_{t=\hat{t}} d\mu
\end{aligned}$$

□

3.5. **Riemann v. Lebesgue.** At this point, we have built a flexible and powerful theory of measure and integration, which resolves all issues that motivated our ongoing quest in the first place. However, as it is (or at least, should be) the case with all new ideas, our new theory is expected to be just as powerful as the old one and generalize it if possible.

In this subsection, we shall first try to understand the relationship between Riemann integral and Lebesgue integral. Then we are going to characterize Riemann integrable functions. By Lebesgue integral, we mean the integral over the measure space $(\mathbb{R}, \mathfrak{L}, \mathbf{m})$ defined in the previous subsection.

A function $f : \mathbb{R} \to \mathbb{R}$ is said to be *Lebesgue integrable* if it is Lebesgue measurable and $\int_{\mathbb{R}} |f| \; d\mathbf{m} < \infty$. In this case, its *Lebesgue integral* is defined to be $\int_{\mathbb{R}} f \; d\mathbf{m}$.[11] One can similarly define Lebesgue integrability a function $f : S \to \mathbb{R}$ with domain $S \subseteq \mathbb{R}$ by simply considering its extension $\hat{f} : \mathbb{R} \to \mathbb{R}$ taking the value 0 on S^c.

Fulfilling our expectations, any function that is (proper) Riemann integrable over a compact interval is Lebesgue integrable over the same interval and its Lebesgue integral is the same as its Riemann integral, which allows us to use the Fundamental Theorem of Calculus to compute $\int_{[a,b]} f \; d\mathbf{m}$ for Riemann integrable functions.

Theorem 30. *Let $f : [a, b] \to \mathbb{R}$ be a bounded function. If f is Riemann integrable over $[a, b]$, then f is Lebesgue integrable over $[a, b]$ and moreover, we have*

$$\int_a^b f(x)dx = \int_{[a,b]} f \; d\mathbf{m}$$

Proof. Throughout the proof, we retain the notation from Section 0.1 for partitions and Riemann sums. Without loss of generality, we shall also assume that the domain of f is $\mathbb{R}$ and $f(x) = 0$ for all $x \notin [a, b]$.

Assume that f is Riemann integrable over $[a, b]$. Then there exists a sequence $(P_k)_{k\in\mathbb{N}}$ of partitions of $[a, b]$ such that $P_0 \subseteq P_1 \subseteq \dots$ and

$$\lim_{k\to\infty} U(f, P_k) = \lim_{k\to\infty} L(f, P_k) = \int_a^b f(x)dx$$

For each $k \in \mathbb{N}$, define the simple functions

$$g_k(x) = \sum_{i=1}^{n_k} \left(\inf_{t\in[t^k_{i-1},t^k_i]} f(t) \right) \cdot \chi_{[t^k_{i-1},t^k_i)}(x) + f(b) \cdot \chi_{\{b\}}(x)$$

and

$$G_k(x) = \sum_{i=1}^{n_k} \left(\sup_{t\in[t^k_{i-1},t^k_i]} f(t) \right) \cdot \chi_{[t^k_{i-1},t^k_i)}(x) + f(b) \cdot \chi_{\{b\}}(x)$$

where $P_k = \{t^k_1, t^k_2, \dots, t^k_{n_k}\}$ and $a = t^k_1 < \dots < t^k_{n_k} = b$. Recalling how the integral of a simple function is computed over $(\mathbb{R}, \mathfrak{L}, \mathbf{m})$, one easily deduces that $\int_{\mathbb{R}} g_k \; d\mathbf{m} = L(f, P_k)$ and $\int_{\mathbb{R}} G_k \; d\mathbf{m} = U(f, P_k)$ for every $k \in \mathbb{N}$.

Set $G(x) = \lim_{k\to\infty} G_k(x)$ and $g(x) = \lim_{k\to\infty} g_k(x)$ for all $x \in \mathbb{R}$. Note that these limits exists for all $x \in \mathbb{R}$, since $g_k(x) \le f(x) \le G_k(x)$ for all $x \in \mathbb{R}$ and $k \in \mathbb{N}$. Consequently, g and G are Lebesgue (indeed, Borel) measurable functions. Moreover, $g(x) \le f(x) \le G(x)$ for all $x \in \mathbb{R}$. Applying the Dominated Convergence

---

[11] We would like to remind the reader the following important subtle point. When we defined the integral of functions $f : \mathbf{X} \to \mathbb{R}$ in $L(\mathbf{X}, \mathcal{M}, \mu)$, the codomain $\mathbb{R}$ of these functions was endowed with its **Borel** structure $\mathcal{B}(\mathbb{R})$. This means that, while talking about Lebesgue integrability of a function $f : \mathbb{R} \to \mathbb{R}$ in $L(\mathbb{R}, \mathfrak{L}, \mathbf{m})$, the domain $\mathbb{R}$ is endowed with its Lebesgue σ-algebra $\mathfrak{L}$ whereas the codomain $\mathbb{R}$ is endowed with its Borel σ-algebra $\mathcal{B}(\mathbb{R})$.

Theorem to the sequences $(g_k)_{k\in\mathbb{N}}$ and $(G_k)_{k\in\mathbb{N}}$ with the dominating integrable function G_0, we obtain that

$$\int_{\mathbb{R}} G\ d\mathbf{m} = \lim_{k\to\infty}\int_{\mathbb{R}} G_k\ d\mathbf{m} = \lim_{k\to\infty} U(f,P_k)$$
$$= \lim_{k\to\infty} L(f,P_k) = \lim_{k\to\infty}\int_{\mathbb{R}} g_k\ d\mathbf{m} = \int_{\mathbb{R}} g\ d\mathbf{m}$$

It now follows from Lemma 5 that $g(x) = G(x)$ holds $\mathbf{m}$-almost everywhere and hence $g = f = G$ holds $\mathbf{m}$-almost everywhere. Since $(\mathbb{R}, \mathfrak{L}, \mathbf{m})$ is complete and g is Lebesgue measurable, Proposition 8 gives us that f is Lebesgue measurable. Since f is bounded, it is Lebesgue integrable over $[a,b]$ and moreover, $g = f$ holding $\mathbf{m}$-almost everywhere implies that

$$\int_{[a,b]} f\ d\mathbf{m} = \int_{\mathbb{R}} f\ d\mathbf{m} = \int_{\mathbb{R}} g\ d\mathbf{m} = \lim_{k\to\infty} L(f,P_k) = \int_a^b f(x)dx$$

□

Unfortunately, the analogue of this theorem may fail for *improper* Riemann integrals. For example, consider the measurable map $f : (0,\infty) \to \mathbb{R}$ given by $f(x) = (-1)^{\lceil x\rceil}\lceil x\rceil^{(-1)}$ for all $x \in (0,\infty)$. A quick calculation shows that

$$\int_0^\infty f(x)dx = -\frac{1}{1} + \frac{1}{2} - \frac{1}{3} + \cdots = \ln(1/2)$$

whereas

$$\int_{(0,\infty)} f^+\ d\mathbf{m} = \frac{1}{2}+\frac{1}{4}+\frac{1}{6}+\cdots = +\infty \text{ and } \int_{(0,\infty)} f^-\ d\mathbf{m} = -\frac{1}{1}-\frac{1}{3}-\frac{1}{5}+\cdots = -\infty$$

showing that f is not Lebesgue integrable over $(0,\infty)$. Although improper Riemann integrals existing does not imply Lebesgue integrability in general, under certain hypotheses, this may be the case. For example, using the Monotone Convergence Theorem, the reader may try to prove that if $f : (a,\infty) \to \mathbb{R}$ is a non-negative bounded function that is Riemann integrable over all compact subintervals of (a,∞) such that $\int_a^\infty f(x)dx < \infty$, then it is Lebesgue integrable over (a,∞) and $\int_a^\infty f(x)dx = \int_{(a,\infty)} f\ d\mathbf{m}$. More generally, it follows from the Dominated Convergence Theorem that, for a (possibly unbounded) interval $I \subseteq \mathbb{R}$ and a map $f : I \to \mathbb{R}$, if f and $|f|$ are proper or improper Riemann integrable, then f is Lebesgue integrable and its Lebesgue integral equals its Riemann integral. See [Bog07, Theorem 2.10.2] for a proof of this fact.

Next shall be characterized Riemann integrable functions over compact intervals. Recall the basic calculus fact that continuous functions over compact intervals are Riemann integrable. With some more effort, one can generalize this fact to functions with finitely and countably many discontinuities. It turns out that even this can be generalized and that Riemann integrable functions are exactly those whose discontinuities form a Lebesgue null set.

Theorem 31 (Lebesgue's criterion for Riemann integrability)**.** *Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Then f is Riemann integrable over $[a, b]$ if and only if $\mathbf{m}(\{x \in [a, b] : f$ is discontinuous at $x\}) = 0$.*

Before we prove this fact, we will introduce some auxiliary notions that are going to be needed in the proof. Let $f : S \to \mathbb{R}$ be a function and $D \subseteq S \subseteq \mathbb{R}$. The *oscillation of f over D* is defined as

$$\omega_f(D) = \sup_{x,y \in D} |f(x) - f(y)|$$

Thus $\omega_f(D)$ measures the width of the thinnest horizontal strip that can contain the graph of f over D. Given $x \in S$, the *oscillation of f at x* is defined as

$$\omega_f(x) = \inf_{\delta > 0} \omega_f(B(x, \delta) \cap dom(f))$$

Intuitively speaking, $\omega_f(x)$ measures "how much discontinuous" f is at x. Indeed, the reader is expected to check that f is continuous at x if and only if $\omega_f(x) = 0$.

Proposition 11. *Let $f : \mathbb{R} \to \mathbb{R}$ be a function and let $\epsilon > 0$. Set*

$$S_{f,\epsilon} = \{x \in \mathbb{R} : \omega_f(x) < \epsilon\}$$

Then $S_{f,\epsilon}$ is open.

Proof. Let $\hat{x} \in S_{f,\epsilon}$. Then $\omega_f(\hat{x}) = \inf_{\delta > 0} \omega_f(B(\hat{x}, \delta)) < \epsilon$ and hence, there exists $\hat{\delta} > 0$ such that $\omega_f(B(\hat{x}, \hat{\delta})) < \epsilon$. We claim that $B(\hat{x}, \hat{\delta}) \subseteq S_{f,\epsilon}$, which would show that $S_{f,\epsilon}$ is open. Let $z \in B(\hat{x}, \hat{\delta})$. Then, for some $\delta > 0$, we have $B(z, \delta) \subseteq B(\hat{x}, \hat{\delta})$ and so $\omega_f(B(z, \delta)) \leq \omega_f(B(\hat{x}, \hat{\delta}))$. Consequently, $\omega_f(z) \leq \omega_f(\hat{x}) < \epsilon$. Thus $z \in S_{f,\epsilon}$ and so $B(\hat{x}, \hat{\delta}) \subseteq S_{f,\epsilon}$. □

We are now ready to prove Theorem 31.

Proof of Theorem 31. Set $D = \{x \in [a, b] : f$ is discontinuous at $x\}$. Assume that f is Riemann integrable over $[a, b]$. We want to prove that $\mathbf{m}(D) = 0$. It is clear that

$$D = \{x \in [a, b] : \omega_f(x) > 0\} = \bigcup_{n \in \mathbb{N}^+} \left\{x \in [a, b] : \omega_f(x) \geq \frac{1}{n}\right\}$$

Thus, in order to prove $\mathbf{m}(D) = 0$, it suffices to prove that

$$\mathbf{m}\left(\left\{x \in [a, b] : \omega_f(x) \geq \frac{1}{n}\right\}\right) = 0$$

for every $n \in \mathbb{N}^+$. Let $n \in \mathbb{N}^+$ and let $\epsilon > 0$. We will construct a covering of D with open intervals whose total length adds up to less than ϵ. Since f is Riemann integrable, there exists a sequence $(P_k)_{k \in \mathbb{N}}$ of partition of $[a, b]$ such that

$$\lim_{k \to \infty} U(f, P_k) = \lim_{k \to \infty} L(f, P_k) = \int_a^b f(x) dx$$

It follows that $U(f,P_k) - L(f,P_k) < \epsilon/2n$ for some $k \in \mathbb{N}$. Say, $P_k = \{x_0, x_1, \ldots, x_m\}$ with $a = x_0 < x_1 < \cdots < x_m = b$. Set

$$I = \left\{ i \in \mathbb{N} : 0 \leq i \leq m-1 \text{ and } [x_i, x_{i+1}] \cap \left\{ x \in [a,b] : \omega_f(x) \geq \frac{1}{n} \right\} \neq \emptyset \right\}$$

Then clearly

$$\left\{ x \in [a,b] : \omega_f(x) \geq \frac{1}{n} \right\} \subseteq \bigcup_{i \in I} [x_i, x_{i+1}] \subseteq \bigcup_{i \in I} \left(x_i - \frac{\epsilon}{4m}, x_{i+1} + \frac{\epsilon}{4m} \right)$$

and moreover,

$$\begin{aligned}
U(f,P_k) - L(f,P_k) &< \epsilon/2n \\
\sum_{i=0}^{m-1} \left(\sup_{x \in [x_i, x_{i+1}]} f(x) - \inf_{x \in [x_i, x_{i+1}]} f(x) \right) (x_{i+1}, x_i) &< \epsilon/2n \\
\sum_{i=0}^{m-1} \omega_f([x_{i+1}, x_i])(x_{i+1} - x_i) &< \epsilon/n \\
\sum_{i \in I} \omega_f([x_{i+1}, x_i])(x_{i+1} - x_i) \leq \sum_{i=0}^{m-1} \omega_f([x_{i+1}, x_i])(x_{i+1} - x_i) &< \epsilon/2n \\
\sum_{i \in I} \frac{1}{n}\ (x_{i+1}, x_i) \leq \sum_{i \in I} \omega_f([x_{i+1}, x_i])(x_{i+1} - x_i) &< \epsilon/2n \\
\sum_{i \in I}\ (x_{i+1} - x_i) &< \epsilon/2
\end{aligned}$$

Therefore, we have that

$$\mathbf{m}\left(\left\{ x \in [a,b] : \omega_f(x) \geq \frac{1}{n} \right\} \right) < \frac{\epsilon}{2} + \frac{m \cdot \epsilon}{2m} = \epsilon$$

As $\epsilon > 0$ was arbitrary, this shows that $\mathbf{m}\left(\left\{x \in [a,b] : \omega_f(x) \geq \frac{1}{n}\right\}\right) = 0$ for all $n \in \mathbb{N}^+$ implying that $\mathbf{m}(D) = 0$.

For the converse direction, assume that $\mathbf{m}(D) = 0$. Let $\epsilon > 0$. Consider the set

$$E = \left\{ x \in [a,b] : \omega_f(x) \geq \frac{\epsilon}{2(b-a)} \right\}$$

Clearly $E \subseteq D$ and so $\mathbf{m}(E) = 0$. It follows that there exists a sequence $(U_i)_{i \in \mathbb{N}}$ of open intervals such that $E \subseteq \bigcup_{i \in \mathbb{N}} U_i$ and

$$\sum_{i=0}^{\infty} |U_i| < \frac{\epsilon}{4K}$$

By Proposition 11, E is closed and, being also bounded, it is compact. This means that $E \subseteq \bigcup_{k=0}^{N} U_{i_k}$ for some natural numbers $i_0 < i_1 < \cdots < i_N$. We may assume without loss of generality that U_{i_k}'s do not intersect for, otherwise, we could take the union of those that are intersecting.

Set $P = \{x \in [a,b] : x \text{ is an endpoint of } U_{i_k} \text{ for some } 0 \leq k \leq N\} \cup \{a,b\}$ Say, $P = \{x_0, x_1, \ldots, x_M\}$ where $a = x_0 < x_1 < \cdots < x_M = b$. Consider the set of indices

$$J = \left\{ i \in \mathbb{N} : 0 \leq i \leq M \text{ and } (x_i, x_{i+1}) \subseteq \bigcup_{i \in \mathbb{N}} U_i \right\}$$

Note that if $i \notin J$, then $(x_i, x_{i+1}) \subseteq E^c$. Also note that, since f is bounded on $[a,b]$, there exists a constant $K > 0$ such that $|f(x)| < K$ for all $x \in [a,b]$. Computing the difference between the corresponding upper and lower Riemann sums, we get

$$\begin{aligned} U(f,P) - L(f,P) &= \sum_{i=0}^{M-1} \left(\sup_{x \in [x_i, x_{i+1}]} f(x) - \inf_{x \in [x_i, x_{i+1}]} f(x) \right) (x_{i+1} - x_i) \\ &\leq \sum_{i \in J} \omega_f([x_{i+1}, x_i])(x_{i+1} - x_i) + \sum_{i \notin J} \omega_f([x_{i+1}, x_i])(x_{i+1} - x_i) \\ &\leq \sum_{i \in J} 2K(x_{i+1} - x_i) + \sum_{i \notin J} \frac{\epsilon}{2(b-a)}(x_{i+1} - x_i) \\ &\leq K \sum_{i \in J} (x_{i+1} - x_i) + \frac{\epsilon}{2(b-a)} \sum_{i \notin J} (x_{i+1} - x_i) \\ &\leq \frac{2K \cdot \epsilon}{4K} + \frac{\epsilon \cdot (b-a)}{2(b-a)} = \epsilon \end{aligned}$$

Since f is a bounded function on $[a,b]$ for which we can make the difference between upper and lower Riemann sums arbitrarily small, f is Riemann integrable over $[a,b]$, which finishes the proof of the theorem. □

3.6. **Modes of convergence.** In this subsection, we are going to analyze the relationship between different types of "convergence" of real-valued function sequences over a measure space. Throughout this subsection, we shall work on a fixed measure space $(\mathbf{X}, \mathcal{M}, \mu)$.

Let $(f_n)_{n \in \mathbb{N}}$ be a sequence of functions with $f_n : \mathbf{X} \to \mathbb{R}$ for all $n \in \mathbb{N}$ and $f : \mathbf{X} \to \mathbb{R}$ be a function. The notion that

$$(f_n)_{n \in \mathbb{N}} \text{ converges to } f$$

written $f_n \to f$ for shorthand, can be interpreted in various useful ways. We shall now list some of these fundamental convergence types. We say that

- $f_n \to f$ pointwise on $\mathbf{X}$ if

$$\text{For every } x \in \mathbf{X} \ \lim_{n \to \infty} f_n(x) = f(x)$$

  In other words, $f_n \to f$ pointwise if

$$\forall x \in \mathbf{X} \ \forall \epsilon \in \mathbb{R}^+ \ \exists k \in \mathbb{N} \ \forall n \geq k \ |f_n(x) - f(x)| < \epsilon$$

- $f_n \to f$ uniformly on $\mathbf{X}$ if

$$\forall \epsilon \in \mathbb{R}^+ \ \exists k \in \mathbb{N} \ \forall x \in \mathbf{X} \ \forall n \geq k \ |f_n(x) - f(x)| < \epsilon$$

- $f_n \to f$ μ-almost everywhere on $\mathbf{X}$ if
$$\mu(\{x \in \mathbf{X} : \lim_{n\to\infty} f_n(x) \neq f(x)\}) = 0$$
In other words, $f_n \to f$ μ-almost everywhere on $\mathbf{X}$ if
$$\exists N \in \mathbb{M}_{null}\ \forall x \in N^c\ \forall \epsilon \in \mathbb{R}^+\ \exists k \in \mathbb{N}\ \forall n \geq k\ |f_n(x) - f(x)| < \epsilon$$
- $f_n \to f$ in L^1 on $\mathbf{X}$ if
$$\lim_{n\to\infty} \int_{\mathbf{X}} |f_n - f|\ d\mu = 0$$
assuming that $f_n \in L^1(\mathbf{X}, \mathcal{M}, \mu)$ for all $n \in \mathbb{N}$ and $f \in L^1(\mathbf{X}, \mathcal{M}, \mu)$.
- $f_n \to f$ in measure on $\mathbf{X}$ if
$$\forall \epsilon \in \mathbb{R}^+\ \lim_{n\to\infty} \mu(\{x \in \mathbf{X} :\ |f_n(x) - f(x)| \geq \epsilon\}) = 0$$
- $f_n \to f$ almost uniformly $\mathbf{X}$ if
$$\forall \epsilon \in \mathbb{R}^+\ \exists M \in \mathcal{M}\ \mu(M) < \epsilon \text{ and } f_n \to f \text{ uniformly on } M^c$$

It is trivial to see that uniform convergence implies both pointwise convergence and almost uniform convergence; and pointwise convergence implies convergence μ-almost everywhere. Observe that uniform convergence and pointwise convergence have nothing to do with the underlying measure space and therefore, is of little significance to us for the purposes of this course. Let us now show the basic implications between the other types of convergence.

Theorem 32. *Let $f_n \in L(\mathbf{X}, \mathcal{M}, \mu)$ for all $n \in \mathbb{N}$ and $f \in L(\mathbf{X}, \mathcal{M}, \mu)$. Then*

a. *If $f_n \to f$ in L^1, then $f_n \to f$ in measure.*
b. *If $f_n \to f$ almost uniformly, then $f_n \to f$ in measure.*
c. *If $f_n \to f$ almost uniformly, then $f_n \to f$ μ-almost everywhere.*

Proof. Let us prove (a). Assume that $f_n \to f$ in L^1. Let $\epsilon > 0$. Set $E_n = \{x \in \mathbf{X} : |f_n(x) - f(x)| \geq \epsilon\}$ for each $n \in \mathbb{N}$. Then each E_n is measurable and moreover,
$$\int_{\mathbf{X}} |f_n - f|\ d\mu \geq \int_{E_n} |f_n - f|\ d\mu \geq \int_{E_n} \epsilon\ d\mu = \epsilon \cdot \mu(E_n) \geq 0$$
As $f_n \to f$ in L^1, we have $\lim_{n\to\infty} \mu(E_n) = 0$, which means that $f_n \to f$ in measure.

To prove (b), assume that $f_n \to f$ almost uniformly. Let $\epsilon \in \mathbb{R}^+$. We wish to show $\lim_{n\to\infty} \mu(\{x \in \mathbf{X} :\ |f_n(x) - f(x)| \geq \epsilon\}) = 0$. As $f_n \to f$ almost uniformly, there exists $M \in \mathcal{M}$ such that $\mu(M) < \epsilon$ and $f_n \to f$ uniformly on M^c. It follows that there exists $k \in \mathbb{N}$ such that for every $n \geq k$ and for every $x \in M^c$, we have $|f_n(x) - f(x)| < \epsilon$. Consequently,
$$\mu(\{x \in \mathbf{X} :\ |f_n(x) - f(x)| \geq \epsilon\}) \leq \mu(M) < \epsilon$$
for all $n \geq k$, which means that $\lim_{n\to\infty} \mu(\{x \in \mathbf{X} :\ |f_n(x) - f(x)| \geq \epsilon\}) = 0$.

To prove (c), assume that $f_n \to f$ almost uniformly. Then, for every $k \in \mathbb{N}^+$, there exists $M_k \in \mathcal{M}$ such that $\mu(M_k) < 1/k$ and $f_n \to f$ uniformly on M_k^c. Set $M = \bigcap_{k\in\mathbb{N}^+} M_k$. Then clearly $M \in \mathcal{M}$ and $\mu(M) = 0$. Let $x \in M^c$. Then $x \in M_k^c$ for some $k \in \mathbb{N}^+$ and hence, by definition, $\lim_{n\to\infty} f_n(x) = f(x)$. Since $\mu(M) = 0$, we have that $f_n \to f$ μ-almost everywhere. ☐

Let us next see why there are no more implications between these types of convergence for arbitrary measure spaces, by providing counterexamples to the remaining implications.

a. Convergence in measure does not imply convergence in L^1.

   Consider the measure space $(\mathbb{R}, \mathfrak{L}, \mathbf{m})$. Let $f_n = n \cdot \chi_{(0,1/n)}$ for all $n \in \mathbb{N}^+$. Then $f_n \to \mathbf{0}$ in measure since, for all $\epsilon \in \mathbb{R}^+$ and $n \in \mathbb{N}$,

$$\mu\left(\{x \in \mathbb{R} : |f_n(x) - \mathbf{0}(x)| \geq \epsilon\}\right) \leq \frac{1}{n}$$

   However, $f_n \nrightarrow \mathbf{0}$ in L^1 since we have $\lim_{n\to\infty} \int_{\mathbb{R}} |f_n - \mathbf{0}| \, d\mathbf{m} = 1$.

b. Convergence in measure does not imply convergence μ-almost everywhere.

   Consider the measure space $(\mathbb{R}, \mathfrak{L}, \mathbf{m})$ and the sequence of functions

$$f_1 = \chi_{[0,1]},\ f_2 = \chi_{\left[\frac{0}{2},\frac{1}{2}\right]},\ f_3 = \chi_{\left[\frac{1}{2},\frac{2}{2}\right]},\ f_4 = \chi_{\left[\frac{0}{4},\frac{1}{4}\right]},\ f_5 = \chi_{\left[\frac{1}{4},\frac{2}{4}\right]},\ \ldots$$

   which, in general, are defined as

$$f_n = \chi_{\left[\frac{j}{2^k},\frac{j+1}{2^k}\right]}$$

   where $n = 2^k + j$ with $0 \leq j < 2^k$. A quick computation shows that

$$m(\{x \in \mathbb{R} : |f_n(x) - \mathbf{0}(x)| > 0\}) < \frac{2}{n}$$

   and hence $f_n \to \mathbf{0}$ in measure. However, for each $x \in [0,1]$, the sequence $(f_n(x))_{n\in\mathbb{N}}$ alternates infinitely many times and hence does not converge. It follows that $f_n \nrightarrow \mathbf{0}$ μ-almost everywhere.

c. Convergence in measure does not imply almost uniform convergence.

   The example in part b works.

d. Convergence μ-almost everywhere does not convergence in L^1.

   Consider the measure space $(\mathbb{R}, \mathfrak{L}, \mathbf{m})$. Let $f_n = \frac{1}{n}\chi_{(0,n)}$ for all $n \in \mathbb{N}$. Then $f_n \to \mathbf{0}$ μ-almost everywhere (indeed, uniformly) but $f_n \nrightarrow \mathbf{0}$ in L^1

since $\lim_{n\to} \int_{\mathbb{R}} |f_n - \mathbf{0}| \, dm = 1$.

e. Convergence μ-almost everywhere does not imply almost uniform convergence.

Consider the measure space $(\mathbb{R}, \mathfrak{L}, \mathbf{m})$. Let $f_n = \chi_{(n,n+1)}$ for all $n \in \mathbb{N}$. Then $f_n \to \mathbf{0}$ μ-almost everywhere (indeed, pointwise) but $f_n \nrightarrow \mathbf{0}$ since one cannot cover the supports of infinitely many f_n's with a set of finite measure.

f. Convergence μ-almost everywhere does not imply convergence in measure.

The example in part e works.

g. Convergence in L^1 does not imply convergence μ-almost everywhere.

The example in part b works.

h. Convergence in L^1 does not imply almost uniform convergence.

The example in part b works.

i. Almost uniform convergence does not imply convergence in L^1.

The example in part d works.

While convergence in measure of a sequence does not imply the μ-almost everywhere of the sequence, it does imply the μ-almost everywhere convergence of a *subsequence*. Before we prove this fact, let us introduce the notion of a sequence being Cauchy in measure. Let $f_n : \mathbf{X} \to \mathbb{R}$ be measurable for all $n \in \mathbb{N}$. We say that $(f_n)_{n\in\mathbb{N}}$ is *Cauchy in measure* if for every $\epsilon > 0$

$$\lim_{m,n\to\infty} \mu(\{x \in \mathbf{X} : |f_m(x) - f_n(x)| \geq \epsilon\}) = 0$$

That is, for every $\epsilon > 0$ and $\hat{\epsilon} > 0$, there exists $k \in \mathbb{N}$ such that for all $m, n \in \mathbb{N}$ with $m, n \geq k$ we have that $\mu(\{x \in \mathbf{X} : |f_m(x) - f_n(x)| \geq \epsilon\}) < \hat{\epsilon}$. We are now ready to prove the main theorem of this subsection.

Theorem 33. *Let $f_n : \mathbf{X} \to \mathbb{R}$ be measurable for all $n \in \mathbb{N}$. Suppose that $(f_n)_{n\in\mathbb{N}}$ is Cauchy in measure. Then*

- *There exist $f \in L(\mathbf{X}, \mathcal{M}, \mu)$ and a subsequence $(f_{n_k})_{k\in\mathbb{N}}$ such that $f_{n_k} \to f$ μ-almost everywhere and $f_n \to f$ in measure.*

- *If $f, g \in L(\mathbf{X}, \mathcal{M}, \mu)$ are such that $f_n \to f$ in measure and $f_n \to g$ in measure, then $f = g$ holds μ-almost everywhere.*

Proof. For each $k \in \mathbb{N}$, letting $\epsilon = \hat{\epsilon} = 2^{-k}$ in the definition of being Cauchy in measure, we can find $n_k \in \mathbb{N}$ such that $\mu(E_k) < 2^{-k}$ where

$$E_k = \{x \in \mathbf{X} : |f_{n_k}(x) - f_{n_{k+1}}(x)| \geq 2^{-k}\}$$

Moreover, we can arrange these n_k's so that $n_1 < n_2 < \dots$. Consider the subsequence $(g_k)_{k\in\mathbb{N}} = (f_{n_k})_{k\in\mathbb{N}}$. Note that, for each $i \in \mathbb{N}$, if $x \notin \bigcup_{k=i}^{\infty} E_k$, then

$$|g_\ell(x) - g_m(x)| \leq |g_\ell(x) - g_{\ell+1}(x)| + \cdots + |g_{m-1}(x) - g_m(x)| \leq \sum_{k=\ell}^{\infty} \frac{1}{2^k} = \frac{1}{2^{\ell-1}}$$

for all $i \leq \ell \leq m$. Thus, for each $i \in \mathbb{N}$ and $x \notin \bigcup_{k=i}^{\infty} E_k$, the sequence $(g_k(x))_{k\in\mathbb{N}}$ is Cauchy. Consider the function $f : \mathbf{X} \to \mathbb{R}$ given by

$$f(x) = \begin{cases} \lim\limits_{k\to\infty} g_k(x) & \text{if } x \notin \bigcap\limits_{i\in\mathbb{N}} \bigcup\limits_{k=i}^{\infty} E_k \\ 0 & \text{otherwise} \end{cases}$$

Then f can easily be checked to be measurable. Clearly $\mu(\bigcup_{k=0}^{\infty} E_k) \leq 2$ and hence

$$\mu\left(\bigcap_{i\in\mathbb{N}} \bigcup_{k=i}^{\infty} E_k\right) = \lim_{i\to\infty} \mu\left(\bigcup_{k=i}^{\infty} E_k\right) \leq \lim_{i\to\infty} \sum_{k=i}^{\infty} \frac{1}{2^k} = \lim_{i\to\infty} \frac{1}{2^{i-1}} = 0$$

By construction, $g_k(x) \to f(x)$ for every $x \in \left(\bigcap_{i\in\mathbb{N}} \bigcup_{k=i}^{\infty} E_k\right)^c$. Therefore $g_k \to f$ μ-almost everywhere. We now check that $g_k \to f$ in measure. Note that the first inequality in the proof actually implies that

$$|g_\ell(x) - f(x)| = |g_\ell(x) - \lim_{m\to\infty} g_m(x)| = \lim_{m\to\infty} |g_\ell(x) - g_m(x)| \leq \frac{1}{2^{\ell-1}}$$

for all $\ell \in \mathbb{N}$ and for all $x \notin \bigcup_{k=\ell}^{\infty} E_k$. Therefore

$$\mu\left(\left\{x \in \mathbf{X} : |g_\ell(x) - f(x)| \geq \frac{1}{2^{\ell-2}}\right\}\right) \leq \mu\left(\bigcup_{k=\ell}^{\infty} E_k\right) \leq \sum_{k=\ell}^{\infty} \frac{1}{2^k} = \frac{1}{2^{\ell-1}}$$

It follows that $g_k \to f$ in measure since, given $\epsilon \in \mathbb{R}^+$, for sufficiently large $\ell \in \mathbb{N}$ we will have $2^{-(\ell-2)} < \epsilon$ and hence

$$\lim_{\ell\to\infty} \mu\left(\{x \in \mathbf{X} : |g_\ell(x) - f(x)| \geq \epsilon\}\right) \leq \lim_{\ell\to\infty} \frac{1}{2^{\ell-1}} = 0$$

We now show that $f_n \to f$ in measure. Given $\epsilon \in \mathbb{R}^+$, by the triangle inequality, we have that

$$\begin{aligned} &\{x \in \mathbf{X} : |f_n(x) - f(x)| \geq \epsilon\} \subseteq \\ &\{x \in \mathbf{X} : |f_n(x) - g_k(x)| \geq \epsilon/2\} \cup \{x \in \mathbf{X} : |g_k(x) - f(x)| \geq \epsilon/2\} \\ &\{x \in \mathbf{X} : |f_n(x) - f_{n_k}(x)| \geq \epsilon/2\} \cup \{x \in \mathbf{X} : |f_{n_k}(x) - f(x)| \geq \epsilon/2\} \end{aligned}$$

for any $n, k \in \mathbb{N}$. Since $(f_n)_{n\in\mathbb{N}}$ is Cauchy in measure and $g_k = f_{n_k} \to f$ in measure, both of the latter sets can be made to have arbitrarily small measure for sufficiently large $n, k \in \mathbb{N}$. It follows that $\lim_{n\to\infty} \mu(\{x \in \mathbf{X} : |f_n(x) - f(x)| \geq \epsilon\}) = 0$ and hence, $f_n \to f$ in measure. This completes the proof of the first part of the theorem.

We now prove the second part via a similar argument. Let $f, g \in L(\mathbf{X}, \mathcal{M}, \mu)$. Assume that $f_n \to f$ in measure and $f_n \to g$ in measure. Then, for any $\epsilon \in \mathbb{R}^+$, we have that

$$\begin{aligned}&\{x \in \mathbf{X} : |f(x) - g(x)| \geq \epsilon\} \subseteq \\ &\{x \in \mathbf{X} : |f(x) - f_n(x)| \geq \epsilon/2\} \cup \{x \in \mathbf{X} : |f_n(x) - g(x)| \geq \epsilon/2\}\end{aligned}$$

for any $n \in \mathbb{N}$. Since the latter sets can be made to have arbitrarily small measure by choosing sufficiently large $n \in \mathbb{N}$, we have that $\mu(\{x \in \mathbf{X} : |f(x) - g(x)| \geq \epsilon\}) = 0$. As this is true for all $\epsilon \in \mathbb{R}^+$, we have that

$$\mu(\{x \in \mathbf{X} : |f(x) - g(x)| > 0\}) = \mu\left(\bigcup_{k\in\mathbb{N}^+} \left\{x \in \mathbf{X} : |f(x) - g(x)| \geq \frac{1}{k}\right\}\right) = 0$$

Thus $f = g$ holds μ-almost everywhere. □

While there are no more implications between these modes of convergence for arbitrary measure spaces, there do exist such implications if one assumes additional hypotheses regarding the measure space. For example, the reader may check that uniform convergence implies convergence in L^1 for finite measure spaces. More importantly, we have Egoroff's theorem, which states that, in finite measure spaces, μ-almost everywhere convergence implies almost uniform convergence.

Theorem 34 (Egoroff's theorem)**.** *Suppose that $(\mathbf{X}, \mathcal{M}, \mu)$ is a finite measure space. Let $f_n : \mathbf{X} \to \mathbb{R}$ be measurable for all $n \in \mathbb{N}$ and $f : \mathbf{X} \to \mathbb{R}$ be measurable. If $f_n \to f$ μ-almost everywhere, then $f_n \to f$ almost uniformly.*

Proof. Assume that $f_n \to f$ μ-almost everywhere. As usual, we will modify these functions on a null set to apply our tools. Set $N = \{x \in \mathbf{X} : \lim_{n\to\infty} f_n(x) = f(x)\}$. Then $\mu(N) = 0$ by assumption. Set $g_n = f_n \cdot \chi_{N^c}$ for each $n \in \mathbb{N}$ and $g = f \cdot \chi_{N^c}$. Clearly we have $g_n \to g$ pointwise. For each $k, n \in \mathbb{N}$, consider the set

$$E_n^k = \bigcup_{m=n}^{\infty} \left\{x \in \mathbf{X} : |g_m(x) - g(x)| \geq \frac{1}{k+1}\right\}$$

Then, for each fixed $k \in \mathbb{N}$, we have $E_0^k \supseteq E_1^k \supseteq \dots$ and, since $g_n \to g$ pointwise, we also have $\bigcap_{n\in\mathbb{N}} E_n^k = \emptyset$. Having $\mu(\mathbf{X}) < \infty$, we can now use Theorem 5.d and obtain $\lim_{n\to\infty} \mu(E_n^k) = 0$.

We next show that $g_n \to g$ almost uniformly. Let $\epsilon \in \mathbb{R}^+$. For each $k \in \mathbb{N}$, using our previous observation that $\lim_{n\to\infty} \mu(E_n^k) = 0$, we choose $n_k \in \mathbb{N}$ such that $\mu(E_{n_k}^k) < \epsilon \cdot 2^{-(k+2)}$. Set $E = \bigcup_{k\in\mathbb{N}} E_{n_k}^k$. Then, clearly $\mu(E) \leq \sum_{k=0}^{\infty} \epsilon \cdot 2^{-(k+2)} < \epsilon$. We claim that $g_n \to g$ uniformly on E^c. Given $\hat{\epsilon} \in \mathbb{R}^+$, choose some $m \in \mathbb{N}$ such that $(m+1)^{-1} \leq \hat{\epsilon}$. Then, for every $x \in E^c$ and for every $n \in \mathbb{N}$ with $n > n_m$, we have that $|g_n(x) - g(x)| < (m+1)^{-1} \leq \hat{\epsilon}$. So $g_n \to g$ uniformly on E^c, which completes the proof that $g_n \to g$ almost uniformly. We leave it the reader to verify that $f_n \to f$ almost uniformly. □

A nice consequence of Egoroff's theorem is Lusin's theorem which states that if $f : [a,b] \to \mathbb{R}$ is measurable, then, for any $\epsilon \in \mathbb{R}^+$, there exists a compact set $K \subseteq [a,b]$ such that $\mathbf{m}([a,b] - K) < \epsilon$ and the restriction $f \restriction K : K \to \mathbb{R}$ is continuous with respect to the subspace topology on K.[12]

A "classical" proof of Lusin's theorem goes as follows. Let $f : [a,b] \to \mathbb{R}$ be measurable. Then there exists a sequence of continuous functions $(f_n)_{n\in\mathbb{N}}$ with $f_n \to f$ in measure.[13] It now follows from Theorem 33 that there exists a subsequence $(f_{n_k})_{k\in\mathbb{N}}$ such that $f_{n_k} \to f$ μ-almost everywhere. Applying Egoroff's theorem, we get that $f_{n_k} \to f$ almost uniformly. This means that, given $\epsilon \in \mathbb{R}^+$, we can find a set $\hat{E} \subseteq [a,b]$ such that $\mathbf{m}(\hat{E}) < \epsilon/2$ and $f_{n_k} \to f$ uniformly on $[a,b] - \hat{E}$. By Theorem 11, we can choose a compact set $K \subseteq [a,b] - \hat{E}$ such that $\mathbf{m}(([a,b]-\hat{E})-K) < \epsilon/2$. Then we have $\mathbf{m}([a,b]-K) < \epsilon$. Since a uniform limit of a sequence of continuous functions is continuous and $f_{n_k} \restriction K \to f \restriction K$ uniformly, we have that $f \restriction K$ is continuous.

While this argument, which employs many tools that we have developed, is perfectly fine, we prefer to provide an elementary proof from [Oxt80, Theorem 8.2] for the following form of Lusin's theorem.

Theorem 35 (Lusin's theorem)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be a measurable function and let $\epsilon \in \mathbb{R}^+$. Then there exists a closed set $K \subseteq \mathbb{R}$ such that $\mathbf{m}(\mathbb{R} - K) < \epsilon$ and $f \restriction K : K \to \mathbb{R}$ is continuous.*

Proof. Let $\{U_i : i \in \mathbb{N}\}$ be a countable base for the topology of $\mathbb{R}$. Then $f^{-1}[U_i]$ is measurable for all $i \in \mathbb{N}$. From a modification of the proof of Theorem 11 follows that there exist a closed set $F_i \subseteq \mathbb{R}$ and an open set $G_i \subseteq \mathbb{R}$ such that $F_i \subseteq f^{-1}[U_i] \subseteq G_i$ and

$$\mathbf{m}(G_i - F_i) < \frac{\epsilon}{2^{i+2}}$$

for all $i \in \mathbb{N}$. Set $E = \bigcup_{i\in\mathbb{N}} G_i - F_i$ and $K = \mathbb{R} - E$. Clearly, K is closed. Moreover, we have that

$$\mathbf{m}(\mathbb{R} - K) = \mathbf{m}(E) \leq \sum_{i=0}^{\infty} \frac{\epsilon}{2^{i+2}} < \epsilon$$

Consider the restriction $f \restriction K : K \to \mathbb{R}$. For any $i \in \mathbb{N}$, we have that

$$(f \restriction K)^{-1}[U_i] = f^{-1}[U_i] \cap K = F_i \cap K = G_i \cap K$$

[12]Note that this is **not** the same as saying the discontinuities of f is contained in K.

[13]For a proof of this non-trivial claim, the reader may check [Bog07, Proposition 2.2.9].

As $\{U_i : i \in \mathbb{N}\}$ is a base for the topology of $\mathbb{R}$, the last equality implies that the inverse images of open subsets of $\mathbb{R}$ are open (and, indeed closed) in the subspace topology of K. Consequently, the map $f \restriction K$ is continuous. $\square$

3.7. **Exercises.** Below you shall find some exercises that you can work on regarding the topics in this section. These exercises are *not* to be handed in as homework assignments.

- Exercises 3, 6, 8, 12, 14, 16, 19, 22, 25, 28, 32, 35, 38, 42 from Chapter 2 of [Fol99].
- Exercises 3, 10 from Chapter 2.1, Exercise 6 from Chapter 2.2, Exercise 3 from Chapter 2.3, Exercises 3, 4, 5 from Chapter 2.5 and Exercises 3, 4 from Chapter 3.1 of [Coh93].
- Let $(\mathbf{X}, \mathcal{M}, \mu)$ be a measure space with $0 < \mu(\mathbf{X}) < \infty$. Consider the function
$$\rho : L(\mathbf{X}, \mathcal{M}, \mu) \times L(\mathbf{X}, \mathcal{M}, \mu) \to [0, \infty)$$
given by
$$\rho(f, g) = \int_X \min\{|f - g|, 1\}\, d\mu$$
Let $f \in L(\mathbf{X}, \mathcal{M}, \mu)$ and let (f_n) be a sequence of functions in $L(\mathbf{X}, \mathcal{M}, \mu)$. Show that if $f_n \to f$ in measure, then for all $\epsilon \in \mathbb{R}^+$ there exists $k \in \mathbb{N}$ such that for all $n \in \mathbb{N}$ with $n \geq k$, we have that $\rho(f_n, f) < \epsilon$.

4. Products

In this section, we shall define the product of multiple measure spaces and investigate the properties of the integral in this product measure space. We begin by constructing the product measure. Throughout this section, fix two measure spaces $(\mathbf{X}, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$.

4.1. **Product measures.** We would first like to define a "reasonable" measure on the product space $(\mathbf{X} \times \mathbf{Y}, \mathcal{M} \otimes \mathcal{N})$. Recall from Section 1.4 that the product σ-algebra $\mathcal{M} \otimes \mathcal{N}$ is generated by the collection
$$\{A \times B : \ A \in \mathcal{M},\ B \in \mathcal{N}\}$$
The sets in this collection are called *measurable rectangles* of $\mathcal{M} \otimes \mathcal{N}$. Our intuition is that the measure of a measurable rectangle $A \times B$ under the to-be-defined product measure should be $\mu(A) \cdot \nu(B)$. In order to construct such a measure, we shall use Carathéodory's extension theorem.

Let $\mathcal{A}$ be the collection of finite disjoint union of measurable rectangles of $\mathcal{M} \otimes \mathcal{N}$. A straightforward but tedious verification shows that $\mathcal{A}$ is an algebra on $\mathbf{X} \times \mathbf{Y}$. To prove this fact, the reader may use [Fol99, Proposition 1.7] together with that

$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$ and $(A \times B)^c = (X \times B^c) \cup (A \times B^c)$. Consider the map $\rho : \mathcal{A} \to [0, \infty]$ given by

$$\rho\left(\bigsqcup_{i=1}^{n}(A_i \times B_i)\right) = \sum_{i=1}^{n} \mu(A_i)\nu(B_i)$$

where $A_i \in \mathcal{M}$ and $B_i \in \mathcal{N}$ for all $i = 1, 2, \ldots, n$. We wish to show that ρ is a premeasure. Since an element of $\mathcal{A}$ can be represented in multiple ways as a disjoint union measurable rectangles, we need to first check that ρ is well-defined. Let $E \in \mathcal{A}$ and suppose that $E = \bigsqcup_{i=1}^{n}(A_i \times B_i) = \bigsqcup_{j=1}^{m}(C_j \times D_j)$ for some $A_i, C_j \in \mathcal{M}$ and $B_i, D_j \in \mathcal{N}$. For each $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$, set $U_{ij} = A_i \cap C_j$ and $V_{ij} = B_i \cap D_j$. Then $U_{ij} \in \mathcal{M}$ and $V_{ij} \in \mathcal{N}$ for all $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$. Moreover, $U_{ij} \times V_{ij} = (A_i \times B_i) \cap (C_j \times D_j)$ and these sets are disjoint. It follows that

$$\begin{aligned}
\rho\left(\bigsqcup_{i=1}^{n}\bigsqcup_{j=1}^{m}(U_{ij} \times V_{ij})\right) &= \sum_{i=1}^{n}\sum_{j=1}^{m} \mu(U_{ij})\nu(V_{ij}) = \sum_{i=1}^{n} \rho\left(\bigsqcup_{j=1}^{m}(U_{ij} \times V_{ij})\right) \\
&= \sum_{i=1}^{n} \rho\left(\bigsqcup_{j=1}^{m}(A_i \times B_i) \cap (C_j \times D_j)\right) \\
&= \sum_{i=1}^{n} \rho\left((A_i \times B_i) \cap \bigsqcup_{j=1}^{m}(C_j \times D_j)\right) \\
&= \sum_{i=1}^{n} \rho\left(A_i \times B_i\right) = \sum_{i=1}^{n} \mu(A_i)\nu(B_i) = \rho\left(\bigsqcup_{i=1}^{n}(A_i \times B_i)\right)
\end{aligned}$$

A similar argument shows that $\rho\left(\bigsqcup_{j=1}^{m}\bigsqcup_{i=1}^{n}(U_{ij} \times V_{ij})\right) = \rho\left(\bigsqcup_{j=1}^{m}(C_j \times D_j)\right)$ which implies that ρ is well-defined. It is clear that $\rho(\emptyset) = 0$. We next check the countable additivity of ρ.

Let $E_1, E_2, \dots \in \mathcal{A}$ be disjoint sets such that $\bigsqcup_{i=1}^{\infty} E_i \in \mathcal{A}$. We wish to show that

$$\rho\left(\bigsqcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \rho(E_i) \tag{1}$$

Note that, by the definition of $\mathcal{A}$, for each $i \in \mathbb{N}^+$, we have that $E_i = \bigsqcup_{k=1}^{n_i}(A_k^i \times B_k^i)$ for disjoint some measurable rectangles $A_k^i \times B_k^i$. It clearly holds that $\rho(E_i) = \sum_{k=1}^{n_i} \rho(A_k^i \times B_k^i)$. Therefore, Equality (1) holds for arbitrary E_i's in $\mathcal{A}$ if it holds for E_i's in $\mathcal{A}$ that are measurable rectangles, because we can replace E_i's with the measurable rectangles $A_k^i \times B_k^i$ in the appropriate order. It follows that, we may assume without loss of generality that E_i's are measurable rectangles. So, suppose that $E_i = M_i \times N_i$ where $M_i \in \mathcal{M}$ and $N_i \in N$.

Assume for the moment that $\bigsqcup_{i=1}^{\infty} E_i$ is a measurable rectangle, say,

$$\bigsqcup_{i=1}^{\infty} E_i = \bigsqcup_{i=1}^{\infty} M_i \times N_i = C \times D$$

for some $C \in \mathcal{M}$ and $D \in \mathcal{N}$. Then we have

$$\begin{aligned}\chi_C(x)\chi_D(y) &= \chi_{C\times D}(x,y) \\ &= \chi_{\bigsqcup_{i=1}^{\infty} M_i \times N_i}(x,y) = \sum_{i=1}^{\infty} \chi_{M_i \times N_i}(x,y) = \sum_{i=1}^{\infty} \chi_{M_i}(x)\chi_{N_i}(y)\end{aligned}$$

Integrating both sides over Y and applying Proposition 9 gives

$$\chi_C(x)\nu(D) = \int_Y \sum_{i=1}^{\infty} \chi_{M_i}(x)\chi_{N_i}(y)\, d\nu = \sum_{i=1}^{\infty} \int_Y \chi_{M_i}(x)\chi_{N_i}(y)\, d\nu = \sum_{i=1}^{\infty} \chi_{M_i}(x)\nu(N_i)$$

Similarly, integrating both sides over X and applying Proposition 9 now gives

$$\begin{aligned}\rho\left(\bigsqcup_{i=1}^{\infty} E_i\right) = \rho(C\times D) = \mu(C)\nu(D) &= \int_X \sum_{i=1}^{\infty} \chi_{M_i}(x)\nu(N_i)\ d\mu \\ &= \sum_{i=1}^{\infty} \int_X \chi_{M_i}(x)\nu(N_i)\ d\mu \\ &= \sum_{i=1}^{\infty} \mu(M_i)\nu(N_i) = \sum_{i=1}^{\infty} \rho(E_i)\end{aligned}$$

Therefore, Equality (1) holds in the case that $\bigsqcup_{i=1}^{\infty} E_i$ is a measurable rectangle. We now explain why this implies that Equality (1) also holds for an arbitrary $\bigsqcup_{i=1}^{\infty} E_i \in \mathcal{A}$.

Observe that, since $\bigsqcup_{i=1}^{\infty} E_i \in \mathcal{A}$, we have $\bigsqcup_{i=1}^{\infty} E_i = \bigsqcup_{k=1}^{n} C_k \times D_k$ for some disjoint measurable rectangles $C_k \times D_k$. Moreover,

$$C_k \times D_k = \bigsqcup_{i=1}^{\infty} (E_i \cap (C_k \times D_k)) = \bigsqcup_{i=1}^{\infty} (M_i \times N_i) \cap (C_k \times D_k) = \bigsqcup_{i=1}^{\infty} (M_i \cap C_k) \times (N_i \cap D_k)$$

and hence the measurable rectangle $C_k \times D_k$ is a countable union of measurable rectangles. We have just proven that, in this case,

$$\rho\left(C_k \times D_k\right) = \sum_{i=1}^{\infty} \rho((M_i \cap C_k) \times (N_i \cap D_k)) = \sum_{i=1}^{\infty} \mu(M_i \cap C_k)\nu(N_i \cap D_k)$$

Therefore, we have

$$\begin{aligned}
\rho\left(\bigsqcup_{i=1}^{\infty} E_i\right) = \rho\left(\bigsqcup_{k=1}^{n} C_k \times D_k\right) &= \sum_{k=1}^{n} \mu(C_k)\nu(D_k) = \sum_{k=1}^{n} \rho(C_k \times D_k) \\
&= \sum_{k=1}^{n} \sum_{i=1}^{\infty} \mu(M_i \cap C_k)\nu(N_i \cap D_k) \\
&= \sum_{i=1}^{\infty} \sum_{k=1}^{n} \mu(M_i \cap C_k)\nu(N_i \cap D_k) \\
&= \sum_{i=1}^{\infty} \rho\left(\bigsqcup_{k=1}^{n} (M_i \cap C_k) \times (N_i \cap D_k)\right) \\
&= \sum_{i=1}^{\infty} \rho\left(\bigsqcup_{k=1}^{n} (M_i \times N_i) \cap (C_k \times D_k)\right) \\
&= \sum_{i=1}^{\infty} \rho\left((M_i \times N_i) \cap \bigsqcup_{k=1}^{n} (C_k \times D_k)\right) \\
&= \sum_{i=1}^{\infty} \rho\left(M_i \times N_i\right) = \sum_{i=1}^{\infty} \rho\left(E_i\right)
\end{aligned}$$

It follows that $\rho : \mathcal{A} \to [0, \infty]$ is a premeasure. By Carathéodory's extension theorem, as $\mathcal{M}(\mathcal{A}) = \mathcal{M} \otimes \mathcal{N}$, there exists a measure

$$\mu \times \nu : \mathcal{M} \otimes \mathcal{N} \to [0, \infty]$$

such that $\mu \times \nu \restriction \mathcal{A} = \rho$. In other words, we have obtained a measure $\mu \times \nu$ on the measurable space $(\mathbf{X} \times \mathbf{Y}, \mathcal{M} \otimes \mathcal{N})$ such that $(\mu \times \nu)(M \times N) = \mu(M)\nu(N)$ for any $M \in \mathcal{M}$ and $N \in \mathcal{N}$.

This measure $\mu \times \nu$ is called a *product measure* of μ and ν. As mentioned at the end of Section 2.3 without a proof, it turns out that, if the premeasure ρ is σ-finite, then the extension given by Carathéodory's theorem is unique. It follows that, if the measure spaces $(\mathbf{X}, \mathcal{M}, \mu)$ and $(\mathbf{Y}, \mathcal{N}, \nu)$ are σ-finite, then so is $(\mathbf{X} \times \mathbf{Y}, \mathcal{A}, \rho)$ and hence, the extension $\mu \times \nu$ is unique, in which case we may talk about *the* product measure of μ and ν.

4.2. **Sections and measurability.** Let $x \in \mathbf{X}$ and $y \in \mathbf{Y}$. Given $E \in \mathcal{M} \otimes \mathcal{N}$, we define the *x-section* of E to be the set

$$E_x = \{y \in \mathbf{Y} : (x, y) \in E\}$$

and the *y-section* of E to be the set

$$E_y = \{x \in \mathbf{X} : (x, y) \in E\}$$

Similarly, given $f : \mathbf{X} \times \mathbf{Y} \to \mathbb{R}$, we define the *x-section* of f to be the function $f_x : \mathbf{Y} \to \mathbb{R}$ given by

$$f_x(y) = f(x, y)$$

for all $y \in \mathbf{Y}$ and the *y-section* of f to be the function $f^y : \mathbf{X} \to \mathbb{R}$ given by

$$f^y(x) = f(x, y)$$

for all $x \in \mathbf{X}$. It turns out that sections of measurable sets and functions are measurable.

Proposition 12. *For all $E \in \mathcal{M} \otimes \mathcal{N}$ and for all $x \in \mathbf{X}$ and for all $y \in Y$, we have $E_x \in \mathcal{N}$ and $E^y \in \mathcal{M}$.*

Proof. Consider the set

$$\mathcal{R} = \{E \in \mathcal{M} \otimes \mathcal{N} : \ \forall x \in \mathbf{X}\ \forall y \in \mathbf{Y}\ (E_x \in \mathcal{N}\ \wedge\ E^y \in \mathcal{M})\}$$

Each measurable rectangle $A \times B \in \mathcal{M} \otimes \mathcal{N}$ is in $\mathcal{R}$ since

$$(A \times B)_x = \begin{cases} \emptyset & \text{if } x \notin A \\ B & \text{if } x \in A \end{cases} \quad \text{and} \quad (A \times B)^y = \begin{cases} \emptyset & \text{if } y \notin B \\ A & \text{if } y \in B \end{cases}$$

We next show that $\mathcal{R}$ is a σ-algebra. Let $E_1, E_2, \dots \in \mathcal{R}$. Then we have that

$$\left(\bigcup_{i=1}^{\infty} E_i\right)_x = \bigcup_{i=1}^{\infty} (E_i)_x \in \mathcal{N} \quad \text{and} \quad \left(\bigcup_{i=1}^{\infty} E_i\right)^y = \bigcup_{i=1}^{\infty} (E_i)^y \in \mathcal{M}$$

for all $x \in \mathbf{X}$ and $y \in \mathbf{Y}$, since $\mathcal{M}$ and $\mathcal{N}$ are σ-algebras. Thus $\bigcup_{i=1}^{\infty} E_i \in \mathcal{R}$. Now let $E \in \mathcal{R}$. Then we have that $(E^c)_x = (E_x)^c \in \mathcal{N}$ and $(E^c)^y = (E^y)^c \in \mathcal{M}$ for all $x \in \mathbf{X}$ and $y \in \mathbf{Y}$, as $\mathcal{M}$ and $\mathcal{N}$ are σ-algebras. It follows that $E^c \in \mathcal{R}$, which completes the proof that $\mathcal{R}$ is a σ-algebra. Recall that the σ-algebra generated by measurable rectangles is $\mathcal{M} \otimes \mathcal{N}$. So, as $\mathcal{R}$ is a σ-algebra containing all measurable rectangles, we have that $\mathcal{M} \otimes \mathcal{N} \subseteq \mathcal{R}$, which proves the claim. □

Proposition 13. *For all measurable maps $f : \mathbf{X} \times \mathbf{Y} \to \overline{\mathbb{R}}$ and for all $x \in \mathbf{X}$ and for all $y \in Y$, we have that $f_x : Y \to \overline{\mathbb{R}}$ and $f^y : \mathbf{X} \to \overline{\mathbb{R}}$ are measurable functions.*

Proof. Let $f : \mathbf{X} \times \mathbf{Y} \to \overline{\mathbb{R}}$ be a measurable map, let $x \in \mathbf{X}$ and $y \in \mathbf{Y}$. Then, for any $B \in \mathcal{B}(\overline{\mathbb{R}})$, by the measurability of f and Proposition 13, we have that

$$(f_x)^{-1}[B] = (f^{-1}[B])_x \in \mathcal{N} \quad \text{and} \quad (f^y)^{-1}[B] = (f^{-1}[B])^y \in \mathcal{M}$$

Thus f_x and f^y are measurable. □

Our main goal in this section is to prove the Fubini-Tonelli theorem. With this in mind, we next introduce a useful auxiliary notion that is going to be used in the proof of the next theorem. Let X be a non-empty set. A collection $\mathcal{C} \subseteq \mathcal{P}(X)$ is said to be a *monotone class* if it is closed under countable increasing unions, i.e. $C_1, C_2, \dots \in \mathcal{C}$ with $C_1 \subseteq C_2 \subseteq \dots$ imply $\bigcup_{i=1}^{\infty} C_i \in \mathcal{C}$; and closed under countable decreasing intersections, i.e. $C_1, C_2, \dots \in \mathcal{C}$ with $C_1 \supseteq C_2 \supseteq \dots$ imply $\bigcap_{i=1}^{\infty} C_i \in \mathcal{C}$. Every σ-algebra is clearly a monotone class, however, not every monotone class needs to be a σ-algebra.

It is straightforward to check that the intersection of monotone classes is a monotone class and hence, one can define the *monotone class generated by* a subset $\mathcal{E} \subseteq \mathcal{P}(X)$ as the monotone class

$$\mathcal{C}(\mathcal{E}) = \bigcap \{\mathcal{C} \subseteq \mathcal{P}(X) : \mathcal{E} \subseteq \mathcal{C} \text{ and } \mathcal{C} \text{ is a monotone class}\}$$

We have the following technical fact which states that the monotone class and the σ-algebra generated by an algebra coincide.

Lemma 6 (The monotone class lemma)**.** *Let $\mathcal{A} \subseteq \mathcal{P}(X)$ be an algebra on a set X. Then $\mathcal{C}(\mathcal{A}) = \mathcal{M}(\mathcal{A})$.*

The proof of the Monotone Class Lemma will be given at the end of this section and we now proceed to prove the main ingredient of Fubini-Tonelli theorem.

Theorem 36. *Suppose that $(\mathbf{X}, \mathcal{M}, \mu)$ and $(\mathbf{Y}, \mathcal{N}, \nu)$ are σ-finite measure spaces. Let $E \in \mathcal{M} \otimes \mathcal{N}$. Then*

a. *The maps $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are measurable.*
b.
$$(\mu \times \nu)(E) = \int_X \nu(E_x)\ d\mu = \int_Y \mu(E^y)\ d\nu$$

Proof. We shall first prove the theorem in the case that $(\mathbf{X}, \mathcal{M}, \mu)$ and $(\mathbf{Y}, \mathcal{N}, \nu)$ are finite measure spaces; and then generalize. Assume that $\mu(X), \nu(Y) < \infty$. Consider the collection

$$\mathcal{R} = \{E \in \mathcal{M} \otimes \mathcal{N} : \ (a) \text{ and } (b) \text{ holds for } E\}$$

Let $A \times B \in \mathcal{M} \otimes \mathcal{N}$ be a measurable rectangle. Then the maps

$$x \mapsto \nu((A \times B)_x) = \nu(B)\chi_A(x) \quad \text{and} \quad y \mapsto \mu((A \times B)^y) = \mu(A)\chi_B(y)$$

are measurable and moreover,

$$\begin{aligned}(\mu \times \nu)(A \times B) = \mu(A)\nu(B) &= \int_X \nu(B)\chi_A(x)\ d\mu = \int_X \nu((A \times B)_x)\ d\mu \\ &= \int_Y \mu(A)\chi_B(y)\ d\nu = \int_Y \mu((A \times B)^y)\ d\nu\end{aligned}$$

Thus all measurable rectangles are in $\mathcal{R}$. Using the linearity of integral and the countable additivity of measures, the reader may easily verify that all finite disjoint unions of measurable rectangles are also in $\mathcal{R}$. Thus $\mathcal{A} \subseteq \mathcal{R}$ where $\mathcal{A}$ is the algebra of finite disjoint union of measurable rectangles. We shall next show that $\mathcal{R}$ is a monotone class.

Let $E_1, E_2, \dots \in \mathcal{C}$ be with $E_1 \subseteq E_2 \subseteq \dots$. Set $E = \bigcup_{i=1}^{\infty} E_i$. For each $n \in \mathbb{N}^+$, define $f_n : \mathbf{X} \to \mathbb{R}$ by $f_n(x) = \nu((E_n)_x)$ for all $x \in \mathbf{X}$ and $g_n : \mathbf{Y} \to \mathbb{R}$ by

$g_n(y) = \mu((E_n)^y)$ for all $y \in \mathbf{Y}$. Then, for every $n \in \mathbb{N}^+$, since $E_n \in \mathcal{R}$, the maps f_n and g_n are measurable. Therefore, the maps

$$f(x) = \nu(E_x) = \nu\left(\left(\bigcup_{n=1}^{\infty} E_n\right)_x\right) = \nu\left(\bigcup_{n=1}^{\infty} (E_n)_x\right) = \lim_{n\to\infty} \nu((E_n)_x) = \lim_{n\to\infty} f_n(x)$$

and

$$g(x) = \nu(E^y) = \nu\left(\left(\bigcup_{n=1}^{\infty} E_n\right)^y\right) = \nu\left(\bigcup_{n=1}^{\infty} (E_n)^y\right) = \lim_{n\to\infty} \nu((E_n)^y) = \lim_{n\to\infty} g_n(x)$$

are also measurable. From the Monotone Convergence Theorem and that E_n's are in $\mathcal{R}$, we deduce that

$$\begin{aligned}
(\mu \times \nu)(E) &= (\mu \times \nu)\left(\bigcup_{n=1}^{\infty} E_n\right) \\
&= \lim_{n\to\infty} (\mu \times \nu)(E_n) \\
&= \lim_{n\to\infty} \int_X \nu((E_n)_x)\ d\mu = \lim_{n\to\infty} \int_Y \mu((E_n)^y)\ d\nu \\
&= \int_X \lim_{n\to\infty} \nu((E_n)_x)\ d\mu = \int_Y \lim_{n\to\infty} \mu((E_n)^y)\ d\nu \\
&= \int_X \nu(E_x)\ d\mu = \int_Y \mu(E^y)\ d\nu
\end{aligned}$$

Therefore $E \in \mathcal{R}$. This shows that $\mathcal{R}$ is closed under countable increasing unions. To show that $\mathcal{R}$ is closed under countable decreasing intersections, one carries out the exact same argument except that at the last step, instead of the Monotone Convergence Theorem, one applies the Dominated Convergence Theorem with the dominating functions $x \mapsto \nu((E_1)_x) \leq \nu(\mathbf{Y}) < \infty$ and $y \mapsto \nu((E_1)^y) \leq \mu(\mathbf{X}) < \infty$ which are integrable since $\mu(X), \nu(Y) < \infty$. Thus $\mathcal{R}$ is a monotone class.

On the one hand, $\mathcal{R}$ is a monotone class with $\mathcal{A} \subseteq \mathcal{R}$ and so $\mathcal{C}(A) \subseteq \mathcal{R}$. On the other hand, $\mathcal{A}$ is an algebra on $\mathbf{X} \times \mathbf{Y}$ and so $\mathcal{C}(\mathcal{A}) = \mathcal{M}(\mathcal{A})$ by the Monotone Class Lemma. Thus $\mathcal{M} \otimes \mathcal{N} = \mathcal{M}(A) = \mathcal{C}(A) \subseteq \mathcal{R}$ which completes the proof of the theorem in the case that $(\mathbf{X}, \mathcal{M}, \mu)$ and $(\mathbf{Y}, \mathcal{N}, \nu)$ are finite measure spaces.

Now suppose that $(\mathbf{X}, \mathcal{M}, \mu)$ and $(\mathbf{Y}, \mathcal{N}, \nu)$ are σ-finite measure spaces, say, we have $\mathbf{X} = \bigcup_{i\in\mathbb{N}} X_i$ and $\mathbf{Y} = \bigcup_{i\in\mathbb{N}} Y_i$ for some non-empty measurable sets $X_0 \subseteq X_1 \subseteq \dots$ and $Y_0 \subseteq Y_1 \subseteq \dots$ with $\mu(X_i), \nu(Y_i) < \infty$. For each $i \in \mathbb{N}$, consider the finite measure spaces

$$(X_i, \mathcal{M}_i, \mu_,i) \text{ and } (Y_i, \mathcal{N}_i, \nu_i)$$

where

$$\mathcal{M} \upharpoonright X_i = \mathcal{M}_i = \{F \cap X_i : F \in \mathcal{M}\} \text{ and } \mathcal{N} \upharpoonright Y_i = \mathcal{N}_i = \{F \cap Y_i : F \in \mathcal{N}\}$$

and $\mu_i(F \cap X_i) = \mu(F \cap X_i)$ and $\nu_i(F \cap Y_i) = \nu(F \cap Y_i)$. It can be checked that $\mathcal{M}_i \otimes \mathcal{N}_i = \{K \cap (X_i \times Y_i) : K \in \mathcal{M} \otimes \mathcal{N}\} = (\mathcal{M} \otimes \mathcal{N}) \upharpoonright (X_i \times Y_i)$.

Let $E \in \mathcal{M} \otimes \mathcal{N}$. We shall show that (a) and (b) hold. Since we have that $\mathbf{X} \times \mathbf{Y} = \bigcup_{i \in \mathbb{N}} X_i \times Y_i$, and $X_0 \times Y_0 \subseteq X_1 \times Y_1 \subseteq \dots$, one may verify that

$$(\mu \times \nu)(E) = \lim_{i \to \infty} (\mu_i \times \nu_i)(E \cap (X_i \times Y_i))$$

Moreover, we have proven that the theorem holds for finite measure spaces and hence, for every $i \in \mathbb{N}$, the product space $(X_i \times Y_i, \mathcal{M}_i \times \mathcal{N}_i, \mu_i \times \nu_i)$ satisfies (a) and (b) for its measurable set $E \cap (X_i \times Y_i) \in \mathcal{M}_i \otimes \mathcal{N}_i$. It is readily checked that $\nu(E_x) = \lim_{i \to \infty} \nu_i((E \cap (X_i \times Y_i))_x)$ and $\mu(E^y) = \lim_{i \to \infty} \mu_i((E \cap (X_i \times Y_i))^y)$. Consequently, the functions $x \mapsto \nu(E_x)$ and $y \mapsto \mu_i(E^y)$ are measurable as they are limits of measurable functions. Thus (a) holds for E. It now follows from the Monotone Convergence Theorem and (b) holding for $E \cap (X_i \times Y_i)$ that

$$\begin{aligned}
&(\mu \times \nu)(E) \\
&= \lim_{i \to \infty} (\mu_i \times \nu_i)(E \cap (X_i \times Y_i)) \\
&= \lim_{i \to \infty} \int_{X_i} \nu_i((E \cap (X_i \times Y_i))_x) \; d\mu = \lim_{i \to \infty} \int_{Y_i} \mu_i((E \cap (X_i \times Y_i))^y) \; d\nu \\
&= \lim_{i \to \infty} \int_X \chi_{X_i}(x) \nu((E \cap (X_i \times Y_i))_x) \; d\mu = \lim_{i \to \infty} \int_Y \chi_{Y_i}(y) \mu((E \cap (X_i \times Y_i))^y) \; d\nu \\
&= \int_X \lim_{i \to \infty} \chi_{X_i}(x) \mu((E \cap (X_i \times Y_i))_x) \; d\mu = \int_Y \lim_{i \to \infty} \chi_{Y_i}(y) \nu((E \cap (X_i \times Y_i))^y) \; d\nu \\
&= \int_X \mu(E_x) \; d\mu = \int_Y \nu(E^y) \; d\nu
\end{aligned}$$

Therefore (b) holds for $E \in \mathcal{M} \otimes \mathcal{N}$, which completes the proof. □

4.3. **The Fubini-Tonelli theorem.** We are now ready to prove the Fubini-Tonelli theorem, which basically says that a "double integral" can be evaluated as two iterated integrals.

Theorem 37 (The Fubini-Tonelli theorem)**.** *Suppose that* $(\mathbf{X}, \mathcal{M}, \mu)$ *and* $(\mathbf{Y}, \mathcal{N}, \nu)$ *are* σ*-finite measure spaces. Then the following hold.*

(Tonelli.) *Let* $f \in L^+(\mathbf{X} \times \mathbf{Y}, \mathcal{M} \otimes \mathcal{N}, \mu \times \nu)$. *Then the maps* $x \mapsto \int_Y f_x(y) \; d\nu$ *and* $y \mapsto \int_X f^y(x) \; d\mu$ *are in* $L^+(\mathbf{X}, \mathcal{M}, \mu)$ *and* $L^+(\mathbf{Y}, \mathcal{N}, \nu)$ *respectively, and moreover,*

$$\int_{\mathbf{X} \times \mathbf{Y}} f(x, y) \; d(\mu \times \nu) = \int_{\mathbf{X}} \left(\int_{\mathbf{Y}} f_x(y) \; d\nu \right) d\mu = \int_{\mathbf{Y}} \left(\int_{\mathbf{X}} f^y(x) \; d\mu \right) d\nu$$

(Fubini.) *Let* $f \in L^1(\mathbf{X} \times \mathbf{Y}, \mathcal{M} \otimes \mathcal{N}, \mu \times \nu)$. *Then we have* $f_x \in L^1(\mathbf{Y}, \mathcal{N}, \nu)$ *holds* μ*-almost everywhere and* $f^y \in L^1(\mathbf{X}, \mathcal{M}, \mu)$ *holds* ν*-almost everywhere. Moreover, setting*

$$g(x) = \begin{cases} \displaystyle\int_{\mathbf{Y}} f_x(y) \; d\nu & \text{if } f_x \in L^1(\mathbf{Y}, \mathcal{N}, \nu) \\ 0 & \text{otherwise} \end{cases}$$

and

$$h(y) = \begin{cases} \int_{\mathbf{X}} f^y(x)\ d\mu & \textit{if } f^y \in L^1(\mathbf{X}, \mathcal{M}, \mu) \\ 0 & \textit{otherwise} \end{cases}$$

we have that

$$\int_{\mathbf{X}\times\mathbf{Y}} f(x,y)\ d(\mu \times \nu) = \int_{\mathbf{X}} g(x)\ d\mu = \int_{\mathbf{Y}} h(y)\ d\nu$$

Proof. We first prove Tonelli's theorem. Observe that Tonelli's theorem holds for characteristic functions of measurable sets by Theorem 36 and hence, it holds for simple functions by the linearity of integral. Let $f \in L^+(\mathbf{X}\times\mathbf{Y}, \mathcal{M}\otimes\mathcal{N}, \mu\times\nu)$. Let $(\varphi_n)_{n\to\mathbb{N}}$ be an increasing sequence of simple functions such that $\varphi_n \to f$ pointwise. A moment's thought reveals that $(\varphi_n)_x \to f_x$ pointwise and $(\varphi_n)^y \to f^y$ pointwise for all $x \in \mathbf{X}$ and $y \in \mathbf{Y}$. Then, the Monotone Convergence Theorem implies that the map $x \mapsto \int_{\mathbf{Y}} f_x(y)\ d\nu$ is the pointwise limit of the maps $x \mapsto \int_{\mathbf{Y}} (\varphi_n)_x(y)\ d\nu$. As Tonelli's theorem holds for simple functions, the latter maps are all measurable and consequently, $x \mapsto \int_{\mathbf{Y}} f_x(y)\ d\nu$ is measurable. By a symmetric argument, one sees that the map $y \mapsto \int_{\mathbf{X}} f^y(x)\ d\mu$ is measurable as well. Finally, applying the Monotone Convergence Theorem and using that Tonelli's theorem holds for simple functions, we obtain that

$$\begin{aligned}
&\int_{\mathbf{X}\times\mathbf{Y}} f(x,y)\ d(\mu\times\nu) \\
&= \int_{\mathbf{X}\times\mathbf{Y}} \lim_{n\to\infty} \varphi_n(x,y)\ d(\mu\times\nu) \\
&= \lim_{n\to\infty} \int_{\mathbf{X}\times\mathbf{Y}} \varphi_n(x,y)\ d(\mu\times\nu) \\
&= \lim_{n\to\infty} \int_{\mathbf{X}} \left(\int_{\mathbf{Y}} (\varphi_n)_x(y)\ d\nu \right) d\mu = \lim_{n\to\infty} \int_{\mathbf{Y}} \left(\int_{\mathbf{X}} (\varphi_n)^y(x)\ d\mu \right) d\nu \\
&= \int_{\mathbf{X}} \left(\int_{\mathbf{Y}} \lim_{n\to\infty} (\varphi_n)_x(y)\ d\nu \right) d\mu = \int_{\mathbf{Y}} \left(\int_{\mathbf{X}} \lim_{n\to\infty} (\varphi_n)^y(x)\ d\mu \right) d\nu \\
&= \int_{\mathbf{X}} \left(\int_{\mathbf{Y}} f_x(y)\ d\nu \right) d\mu = \int_{\mathbf{Y}} \left(\int_{\mathbf{X}} f^y(x)\ d\mu \right) d\nu
\end{aligned}$$

This establishes Tonelli's theorem. Before we proceed to prove Fubini's theorem, the reader should recall that, since a non-negative integrable function cannot take the value $+\infty$ on a set of positive measure, the last equality we obtained implies that if $f \in L^+(\mathbf{X}\times\mathbf{Y}, \mathcal{M}\otimes\mathcal{N}, \mu\times\nu)$ and $\int_{\mathbf{X}\times\mathbf{Y}} f(x,y)\ d(\mu\times\nu) < \infty$, then $\int_{\mathbf{Y}} f_x(y)\ d\nu < \infty$ holds μ-almost everywhere and $\int_{\mathbf{X}} f^y(x)\ d\mu < \infty$ holds ν-almost everywhere.

We are now ready to prove Fubini's theorem. Let $f \in L^1(\mathbf{X}\times\mathbf{Y}, \mathcal{M}\otimes\mathcal{N}, \mu\times\nu)$. Consider the maps f^+ and f^- which are in $L^+(\mathbf{X}\times\mathbf{Y}, \mathcal{M}\otimes\mathcal{N}, \mu\times\nu)$. As f is integrable, $\int_{\mathbf{X}\times\mathbf{Y}} f^+\ d(\mu\times\nu) < \infty$ and $\int_{\mathbf{X}\times\mathbf{Y}} f^-\ d(\mu\times\nu) < \infty$ by definition. Thus, by our previous observation, we have that

- $\int_{\mathbf{Y}}(f^+)_x(y)\ d\nu < \infty$ and $\int_{\mathbf{Y}}(f^-)_x(y)\ d\nu < \infty$ hold μ-almost everywhere.
- $\int_{\mathbf{X}}(f^+)^y(x)\ d\mu < \infty$ and $\int_{\mathbf{X}}(f^-)^y(x)\ d\mu < \infty$ hold ν-almost everywhere.

Clearly $(f^+)_x = (f_x)^+$, $(f_x)^- = (f^-)_x$ and $(f^y)^+ = (f^+)^y$, $(f^y)^- = (f^-)^y$, from which the first claim in Fubini's theorem follows.

For the second claim, choose a measurable set $K \subseteq \mathbf{X}$ such that K^c is μ-null and $\int_{\mathbf{Y}}(f^+)_x(y)\ d\nu < \infty$ and $\int_{\mathbf{Y}}(f^-)_x(y)\ d\nu < \infty$ for all $x \in K$. Similarly, choose a measurable set $L \subseteq \mathbf{Y}$ such that L^c is ν-null and $\int_{\mathbf{X}}(f^+)^y(x)\ d\mu < \infty$ and $\int_{\mathbf{X}}(f^-)^y(x)\ d\mu < \infty$ for all $y \in L$. Applying Tonelli's theorem to f^+ and f^-, we obtain that

$$\begin{aligned}\int_{\mathbf{X}\times\mathbf{Y}} f^+(x,y)\ d(\mu\times\nu) &= \int_{\mathbf{X}}\left(\int_{\mathbf{Y}}(f^+)_x(y)\ d\nu\right) d\mu = \int_{\mathbf{Y}}\left(\int_{\mathbf{X}}(f^+)^y(x)\ d\mu\right) d\nu \\ &= \int_{K}\left(\int_{\mathbf{Y}}(f^+)_x(y)\ d\nu\right) d\mu = \int_{L}\left(\int_{\mathbf{X}}(f^+)^y(x)\ d\mu\right) d\nu\end{aligned}$$

and

$$\begin{aligned}\int_{\mathbf{X}\times\mathbf{Y}} f^-(x,y)\ d(\mu\times\nu) &= \int_{\mathbf{X}}\left(\int_{\mathbf{Y}}(f^-)_x(y)\ d\nu\right) d\mu = \int_{\mathbf{Y}}\left(\int_{\mathbf{X}}(f^-)^y(x)\ d\mu\right) d\nu \\ &= \int_{K}\left(\int_{\mathbf{Y}}(f^-)_x(y)\ d\nu\right) d\mu = \int_{L}\left(\int_{\mathbf{X}}(f^-)^y(x)\ d\mu\right) d\nu\end{aligned}$$

Subtracting the second lines of these equalities side by side gives

$$\begin{aligned}\int_{\mathbf{X}\times\mathbf{Y}} f(x,y)\ d(\mu\times\nu) &= \int_{K}\left(\int_{\mathbf{Y}} f_x(y)\ d\nu\right) d\mu = \int_{L}\left(\int_{\mathbf{X}} f^y(x)\ d\mu\right) d\nu \\ &= \int_K g(x)d\mu = \int_L h(y)d\nu \\ &= \int_{\mathbf{X}} g(x)d\mu = \int_{\mathbf{Y}} h(y)d\nu\end{aligned}$$

This completes the proof of Fubini's theorem. □

We shall next see why the hypotheses of the Fubini-Tonelli theorem are needed by providing counter-examples to cases where they are dropped.

- The σ-finiteness assumption in Tonelli's theorem is needed.

  Consider the measure spaces $([0,1],\mathcal{B}([0,1]),\mathbf{m})$ and $([0,1],\mathcal{B}([0,1]),\eta)$ where η is the counting measure. Then $([0,1],\mathcal{B}([0,1]),\eta)$ is not σ-finite. The set $D = \{(x,x) \in [0,1]\times[0,1] : x \in [0,1]\}$ is clearly closed and hence, is in $\mathcal{B}([0,1]\times[0,1]) = \mathcal{B}([0,1]\otimes\mathcal{B}([0,1])$. It follows that $\chi_D : [0,1]\times[0,1] \to \mathbb{R}$ is measurable. On the other hand, the reader can check that

$$\int_{[0,1]}\left(\int_{[0,1]} \chi_D(x,y)\ d\mathbf{m}(x)\right) d\eta(y) = 0$$

$$\int_{[0,1]} \left(\int_{[0,1]} \chi_D(x,y) \ d\eta(y) \right) d\mathbf{m}(x) = 1$$
$$\int_{[0,1]\times[0,1]} \chi_D(x,y) \ d(\mathbf{m} \times \eta) = \infty$$

Thus, Tonelli's theorem does not hold for χ_D.

- The integrability condition in Fubini's theorem is needed.

Consider the measure space $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$ where μ is the counting measure. Then $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$ is clearly σ-finite as $\mathbb{N} = \bigcup_{n\in\mathbb{N}}\{n\}$. Consider the map $f : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ given by

$$f(m,n) = \begin{cases} 1 & \text{if } m = n \\ -1 & \text{n=m+1} \\ 0 & \text{otherwise} \end{cases}$$

It is easily seen that f is measurable (and indeed, is a simple function.) On the other hand, we have that

$$\int_{\{0,1,,\dots,k\}\times\{0,1,\dots,k\}} |f| \ d(\mu\times\mu) = 2k+1$$

and hence, taking the limit as $k \to \infty$, the Monotone Convergence Theorem implies that

$$\int_{\mathbb{N}\times\mathbb{N}} |f| \ d(\mu\times\mu) = \infty$$

Thus f is not integrable. On the one hand, we have

$$\int_{\mathbb{N}} \left(\int_{\mathbb{N}} f(m,n) \ d\mu(n) \right) d\mu(m) = \int_{\mathbb{N}} \left(\sum_{n=0}^{\infty} f(m,n) \right) d\mu(m)$$
$$= \int_{\mathbb{N}} (f(m,m) + f(m,m+1)) \ d\mu(m) = \int_{\mathbb{N}} 0 \ d\mu(m) = 0$$

On the other hand, we also have

$$\int_{\mathbb{N}} \left(\int_{\mathbb{N}} f(m,n) \ d\mu(m) \right) d\mu(n) =$$
$$\int_{\{0\}} \left(\int_{\mathbb{N}} f(m,n) \ d\mu(m) \right) d\mu(n) + \int_{\mathbb{N}-\{0\}} \left(\int_{\mathbb{N}} f(m,n) \ d\mu(m) \right) d\mu(n) =$$
$$\int_{\{0\}} \left(\sum_{m=0}^{\infty} f(m,n) \right) d\mu(n) + \int_{\mathbb{N}-\{0\}} \left(\sum_{m=0}^{\infty} f(m,n) \right) d\mu(n) =$$
$$\left(\sum_{m=0}^{\infty} f(m,0) \right) \cdot \mu(\{0\}) + \int_{\mathbb{N}-\{0\}} (f(n,n) + f(n-1,n)) \ d\mu(n) = 1 + 0 = 1$$

Thus, Fubini's theorem does not hold for f.

- The measurability condition is needed in Tonelli's theorem.

  Consider the measure space $(\omega_1, \mathcal{M}, \eta)$ where ω_1 is the first uncountable ordinal and
  $$\mathcal{M} = \{A \subseteq \omega_1 : A \text{ or } A^c \text{ is countable}\}$$
  and $\eta : \mathcal{M} \to [0,1]$ is the measure given by
  $$\eta(A) = \begin{cases} 1 & \text{if } A^c \text{ is countable} \\ 0 & \text{if } A \text{ is countable} \end{cases}$$
  Then $(\omega_1, \mathcal{M}, \eta)$ is a finite measure space. Let $\prec$ be the usual ordering on ω_1 and consider the set $E = \{(x,y) \in \omega_1 \times \omega_1 : x \prec y\}$. Then E is not in $\mathcal{M} \otimes \mathcal{M}$ and hence, $\chi_E : \omega_1 \times \omega_1 \to \mathbb{R}$ is not measurable. On the other hand,
  $$\int_{\omega_1} \left(\int_{\omega_1} \chi_E(x,y)\ d\eta(x) \right) d\eta(y) = \int_{\omega_1} 1\ d\eta(y) = \eta(\omega_1) = 1$$
  as $\eta(\{y \in \omega_1 : x \prec y\}) = 1$ for each $x \in \omega_1$ and moreover,
  $$\int_{\omega_1} \left(\int_{\omega_1} \chi_E(x,y)\ d\eta(y) \right) d\eta(x) = \int_{\omega_1} 0\ d\eta(y) = 0$$
  as $\eta(\{x \in \omega_1 : x \prec y\}) = 0$ for each $y \in \omega_1$. Thus, the second equality in Tonelli's theorem does not hold for χ_D.

Before we conclude this section, as promised earlier, we are going to prove the Monotone Class Lemma which was one of the ingredients of the proof of the Fubini-Tonelli theorem for characteristic functions of measurable maps.

Proof of Lemma 6. Since every σ-algebra is a monotone class, $\mathcal{M}(\mathcal{A})$ is a monotone class containing $\mathcal{A}$ as a subset and hence, $\mathcal{C}(\mathcal{A}) \subseteq \mathcal{M}(\mathcal{A})$.

To prove $\mathcal{M}(\mathcal{A}) \subseteq \mathcal{C}(\mathcal{A})$, as $\mathcal{A} \subseteq \mathcal{C}(\mathcal{A})$, it suffices to show that $\mathcal{C}(\mathcal{A})$ is a σ-algebra. On the other hand, $\mathcal{C}(\mathcal{A})$ being a monotone class, it is closed under countable increasing unions. Recall that any algebra which is closed under countable increasing unions is automatically closed under countable unions and hence, is a σ-algebra. Thus, it suffices to show that $\mathcal{C}(\mathcal{A})$ is an algebra. As $\mathcal{A} \subseteq \mathcal{C}(\mathcal{A})$, we have that $X \in \mathcal{C}(\mathcal{A})$. It follows that, in order to show that $\mathcal{C}(\mathcal{A})$ is an algebra, it is enough to show that

(†) For all $E, F \in \mathcal{C}(\mathcal{A})$ we have $E - F, F - E, E \cap F \in \mathcal{C}(\mathcal{A})$.

For each $E \in \mathcal{C}(\mathcal{A})$, consider the collection
$$\mathcal{D}(E) = \{F \in \mathcal{C}(\mathcal{A}) :\ E - F, F - E, E \cap F \in \mathcal{C}(\mathcal{A})\}$$

Observe that (†) holds if and only if $\mathcal{C}(\mathcal{A}) \subseteq \mathcal{D}(E)$ for every $E \in \mathcal{C}(\mathcal{A})$. Thus, if we can show that $\mathcal{A} \subseteq \mathcal{D}(E)$ and $\mathcal{D}(E)$ is a monotone class for every $E \in \mathcal{C}(\mathcal{A})$, then, by the definition of $\mathcal{C}(\mathcal{A})$, we will have that $\mathcal{C}(\mathcal{A}) \subseteq \mathcal{D}(E)$ for every $E \in \mathcal{C}(\mathcal{A})$, which would imply (†).

Let $E \in \mathcal{C}(\mathcal{A})$. Let $F_1, F_2, \dots \in \mathcal{D}(E)$ be such that $F_1 \subseteq F_2 \subseteq \dots$. Then we have that

$$E - \bigcup_{n=1}^{\infty} F_n = E \cap \left(\bigcap_{n=1}^{\infty} F_n^c \right) = \bigcap_{n=1}^{\infty} (E \cap F_n^c) \in \mathcal{C}(\mathcal{A})$$

The last claim holds because, by the assumption that $F_n \in \mathcal{D}(E)$, the $E \cap F_n^c$'s are in the monotone class $\mathcal{C}(A)$ which form a decreasing sequence. By a similar reasoning, we have that

$$\left(\bigcup_{n=1}^{\infty} F_n \right) - E = \left(\bigcup_{n=1}^{\infty} F_n \right) \cap E^c = \bigcup_{n=1}^{\infty} (F_n \cap E^c) \in \mathcal{C}(\mathcal{A})$$

and that

$$E \cap \left(\bigcup_{n=1}^{\infty} F_n \right) = \bigcup_{n=1}^{\infty} (E \cap F_n) \in \mathcal{C}(\mathcal{A})$$

It follows that $\bigcup_{n=1}^{\infty} F_n \in \mathcal{D}(E)$. Thus $\mathcal{D}(E)$ is closed under countable increasing unions. By similar arguments, one can also show that if $F_1, F_2, \dots \in \mathcal{D}(E)$ are such that $F_1 \supseteq F_2 \supseteq \dots$, then $\bigcap_{n=1}^{\infty} F_n \in \mathcal{D}(E)$. It follows that $\mathcal{D}(E)$ is a monotone class for every $E \in \mathcal{C}(\mathcal{A})$

We are now ready to prove that $\mathcal{A} \subseteq \mathcal{D}(E)$ for every $E \in \mathcal{C}(\mathcal{A})$. Let $E \in \mathcal{C}(\mathcal{A})$ and let $F \in \mathcal{A}$. As $\mathcal{A}$ is an algebra and $\mathcal{A} \subseteq \mathcal{C}(\mathcal{A})$, we have that $\mathcal{A} \subseteq \mathcal{D}(F)$. But we have shown that $\mathcal{D}(F)$ is a monotone class and consequently, we have $\mathcal{C}(\mathcal{A}) \subseteq \mathcal{D}(F)$. So $E \in \mathcal{D}(F)$. But the definition of $\mathcal{D}(F)$ is symmetrical in E and F and hence, $F \in \mathcal{D}(E)$. Thus $\mathcal{A} \subseteq \mathcal{D}(E)$.

This completes the proof that $\mathcal{A} \subseteq \mathcal{D}(E)$ and $\mathcal{D}(E)$ is a monotone class for every $E \in \mathcal{C}(\mathcal{A})$ which implies (†) which, in turn, implies that $\mathcal{C}(A)$ is an algebra and hence, is a σ-algebra. Therefore $\mathcal{C}(A) \subseteq \mathcal{M}(\mathcal{A})$. □

The measure space $(\mathbf{X} \times \mathbf{Y}, \mathcal{M} \otimes \mathcal{N}, \mu \times \nu)$ is not complete in general. For practical purposes, instead of this measure space, one may need to work in its completion $(\mathbf{X} \times \mathbf{Y}, \overline{\mathcal{M} \otimes \mathcal{N}}, \overline{\mu \times \nu})$. It turns out that a version of the Fubini-Tonelli theorem holds for this complete measure space. We refer the reader to [Fol99, Theorem 2.39] for a statement of this theorem, proof of which is left as an exercise.

4.4. **Reading assignment and exercises.** The Fubini-Tonelli theorem is one of many results regarding multiple integration that the reader is probably already familiar with in the Riemann case, which extend to the Lebesgue case. For example, under appropriate modification, the usual change of coordinates formula for multiple integration also holds for the n-dimensional Lebesgue integral. The reader may (and indeed, should) read more about the n-dimensional Lebesgue integral on

$\mathbb{R}^n$ in [Fol99, Section 2.6]. We will not be able to cover these topics due to time limitations.

Below you shall find some exercises that you can work on regarding the topics in this section. These exercises are *not* to be handed in as homework assignments.

- Exercises 46, 49, 50, 52 from Chapter 2 of [Fol99].
- Exercise 5 from Chapter 5.1 and Exercise 2 from Chapter 5.2 of [Coh93].
- Consider the completion $(\mathbb{R}\times\mathbb{R}, \overline{\mathcal{B}(\mathbb{R})\otimes\mathcal{B}(\mathbb{R})}, \overline{\mathbf{m}\otimes\mathbf{m}})$ of the product space $(\mathbb{R}\times\mathbb{R}, \mathcal{B}(\mathbb{R})\otimes\mathcal{B}(\mathbb{R}), \mathbf{m}\otimes\mathbf{m})$. Show that there exists a measurable set $K \in \overline{\mathcal{B}(\mathbb{R})\otimes\mathcal{B}(\mathbb{R})}$ such that $K \notin \mathcal{B}(\mathbb{R})\otimes\mathcal{B}(\mathbb{R})$ and $(\mathbf{m}\times\mathbf{m})(K) = 0$ and $K_x \in \mathcal{B}(\mathbb{R})$ for every $x \in \mathbb{R}$.

5. Decomposition and differentiation

One may generalize the notion of a measure on a measurable space so that measures to take negative (and even, complex) values. In this section, we shall make a gentle introduction to the theory of signed measures. However, in no way do we claim to provide a comprehensive treatment of the theory or its applications. We refer the reader to [Fol99, Chapter 3] and [Coh93, Chapter 4] for a more detailed treatment.

5.1. **Signed measures and their decomposition.** Let $(\mathbf{X}, \mathcal{M})$ be a measurable space. A *signed (σ-additive) measure* on the measurable space $(\mathbf{X}, \mathcal{M})$ is a map $\mu : \mathcal{M} \to \overline{\mathbb{R}}$ such that

- $\mu(\emptyset) = 0$ and
- If $A_1, A_2, \dots \in \mathcal{M}$ are disjoint, then $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$.

To avoid ambiguity, we will sometimes refer to (usual) measures as positive measures for their images are contained in $[0, \infty]$. In a similar fashion, a signed measure whose image is contained in $[-\infty, 0]$ may be referred to as a *negative measure*. Given a signed measure μ on $(\mathbf{X}, \mathcal{M})$, the triple $(\mathbf{X}, \mathcal{M}, \mu)$ will be called a *signed measure space*.

Before we proceed, we would like to make two very important remarks regarding this definition. Let μ be a signed measure on $(\mathbf{X}, \mathcal{M})$.

- Suppose that $\sum_{i=1}^{\infty} \mu(A_i)$ converge conditionally for some $A_1, A_2, \dots \in \mathcal{M}$. Then, by the Riemann rearrangement theorem, there exists a permutation $\varphi : \mathbb{N}^+ \to \mathbb{N}^+$ such that
$$\sum_{i=1}^{\infty} \mu(A_i) \neq \sum_{i=1}^{\infty} \mu\left(A_{\varphi(i)}\right) = \mu\left(\bigcup_{i=1}^{\infty} A_{\varphi(i)}\right) = \mu\left(\bigcup_{i=1}^{\infty} A_i\right)$$
which is a contradiction. Thus, the infinite series in the σ-additivity condition for a signed measure converges absolutely, if it converges at all.

- Suppose that $\mu(A) = +\infty$ (respectively, $-\infty$) for some $A \in \mathcal{M}$. Then, since we have
$$\mu(\mathbf{X}) = \mu(A \sqcup A^c) = \mu(A) + \mu(A^c)$$
by σ-additivity, in order for $\mu(\mathbf{X})$ to be defined at all, we must have that $\mu(A^c) \neq -\infty$ (respectively, $+\infty$.) Hence $\mu(\mathbf{X}) = +\infty$ (respectively, $-\infty$.) It follows that μ cannot take the values $+\infty$ and $-\infty$ at the same time. In other words, the codomain of a signed measure is actually $(-\infty, +\infty]$ or $[-\infty, \infty)$.

What are examples of signed measures? Let ν and η be measures on the measurable space $(\mathbf{X}, \mathcal{M})$. Then one can check that $\mu = \nu - \eta$ is a signed measure on $(\mathbf{X}, \mathcal{M})$. Let $f : \mathbf{X} \to \mathbb{R}$ be a measurable map such that $\int_{\mathbf{X}} f^+ \, d\nu$ or $\int_{\mathbf{X}} f^- \, d\nu$ is finite. Then the map ξ defined by $\xi(E) = \int_E f \, d\nu$ for all $E \in \mathcal{M}$ can be checked to be a signed measure on $(\mathbf{X}, \mathcal{M})$. We shall soon see that any signed measure is indeed of these forms.

Before we proceed, let us point out some important difference between (positive) measures and signed measures. A signed measure need not be monotone in the sense that $A \subseteq B$ does not imply $\mu(A) \leq \mu(B)$ in general. For example, consider the signed measure given by $\mu(E) = \int_E x \, d\mathbf{m}$ on $([-1, 1], \mathcal{B}([-1, 1]))$. Then
$$\mu([0, 1]) = 1/2 > 0 = \mu([-1, 1])$$
Therefore Theorem 5.a fails for signed measures. Depending on Theorem 5.a, Theorem 5.b also fails for signed measures. On the other hand, one may check that the proofs of part c and d of Theorem 5 goes through for signed measures and so, we have the following.

Theorem 38. *Let $(\mathbf{X}, \mathcal{M}, \mu)$ be a signed measure space. Then*

a. *If $A_1, A_2, \dots \in \mathcal{M}$ and $A_i \subseteq A_{i+1}$ for every $i \in \mathbb{N}^+$, then*
$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{i \to \infty} \mu(A_i)$$

b. *If $A_1, A_2, \dots \in \mathcal{M}$ and $A_i \supseteq A_{i+1}$ for every $i \in \mathbb{N}^+$ and $-\infty < \mu(A_1) < \infty$, then*
$$\mu\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{i \to \infty} \mu(A_i)$$

Proof. Imitate the proofs of part c and d of Theorem 5. □

Next will be introduced the notions of positive, negative and null sets for a signed measure. Let $(\mathbf{X}, \mathcal{M}, \mu)$ be a signed measure space. A set $E \in \mathcal{M}$ is said to be

- *positive* if $\mu(F) \geq 0$ for every $F \subseteq E$ with $F \in \mathcal{M}$,
- *negative* if $\mu(F) \leq 0$ for every $F \subseteq E$ with $F \in \mathcal{M}$,
- *null* if $\mu(F) = 0$ for every $F \subseteq E$ with $F \in \mathcal{M}$.

It is trivially seen from the definition that any measurable subset of a set that is positive (respectively, negative, null) is also positive (respectively, negative, null.)

Lemma 7. *Let* $(\mathbf{X}, \mathcal{M}, \mu)$ *be a signed measure space and let* $E_1, E_2, \cdots \in \mathcal{M}$ *be positive (respective, negative, null.) Then* $\bigcup_{i=1}^{\infty} E_i$ *is positive (respective, negative, null.)*

Proof. Set $F_1 = E_1$ and $F_{n+1} = E_{n+1} - \bigcup_{i=1}^{n} E_i$ for all $n \in \mathbb{N}^+$. Then clearly $F_n \subseteq E_n$ and so, F_n is positive (respectively, negative, null) for all $n \in \mathbb{N}^+$. Let $F \subseteq \bigcup_{i=1}^{\infty} F_i = \bigcup_{i=1}^{\infty} E_i$ be such that $F \in \mathcal{M}$. Then we have that $F \cap F_n \in \mathcal{M}$ for all $n \in \mathbb{N}^+$ and so, by the positivity (respectively, negativity, nullity) of F_n's, we obtain that

$$\mu(F) = \mu\left(\bigcup_{i=1}^{\infty}(F \cap F_i)\right) = \sum_{i=1}^{\infty} \mu(F \cap F_i) \geq 0$$

(respectively, $\mu(F) \leq 0$, $\mu(F) = 0$.) □

Our next goal is to prove that any signed measure space can be decomposed into a positive set and a negative set, essentially in a unique manner. Towards this goal, we shall prove the following lemma which is also important on its own.

Lemma 8. *Let* $(\mathbf{X}, \mathcal{M}, \mu)$ *be a signed measure space and let* $A \in \mathcal{M}$ *be such that* $-\infty < \mu(A) < 0$. *Then there exists a negative set* $B \subseteq A$ *such that* $\mu(B) \leq \mu(A)$.

Proof. Set $A_0 = \emptyset$ and $\delta_0 = 0$. For each $n \in \mathbb{N}^+$, we can recursively choose some $A_n \in \mathcal{M}$ such that

$$\mu(A_n) \geq \min\left\{\frac{\delta_n}{2}, 1\right\}$$

where

$$\delta_n = \sup\left\{\mu(E) : \ E \in \mathcal{M}, \ E \subseteq A - \left(\bigcup_{i=0}^{n-1} A_i\right)\right\}$$

Note that the collections of E's in the definition of δ_n's are non-empty for they contain $\emptyset$ and consequently, $\delta_n \geq 0$ for all $n \in \mathbb{N}$. Moreover, by construction, A_n's are disjoint.

Set $A_\infty = \bigcup_{i \in \mathbb{N}} A_i$ and $B = A - A_\infty$. We claim that B is as expected. Clearly $A_\infty \in \mathcal{M}$ and so $B \in \mathcal{M}$. Moreover, as $\mu(A_n) \geq 0$ for all $n \in \mathbb{N}$, we get that

$$\mu(A) = \mu(A_\infty) + \mu(B) = \mu\left(\bigsqcup_{n \in \mathbb{N}} A_n\right) + \mu(N) = \left(\sum_{i \in \mathbb{N}} \mu(A_n)\right) + \mu(B) \geq \mu(B)$$

It remains to check that B is a negative set. Let $E \subseteq B$ be measurable. Then, by the choice of δ_n's, because $E \subseteq A - \bigcup_{i \in \mathbb{N}} A_i$, we have that $\mu(E) \leq \delta_n$ for all $n \in \mathbb{N}$. Thus it suffices to prove $\lim_{n\to\infty} \delta_n = 0$. Observe that $\mu(A)$ being finite implies that $\mu(A_\infty)$ is finite as $\mu(A) = \mu(A_\infty) + \mu(B)$. It now follows from this observation and $\mu(A_\infty) = \sum_{i \in \mathbb{N}} \mu(A_n)$ that $\lim_{n\to\infty} \mu(A_n) = 0$. Consequently, $\lim_{n\to\infty} \delta_n = 0$ and so, $\mu(E) \leq 0$. Therefore B is a negative set. □

We are now ready to prove the first main result of section.

Theorem 39 (Hahn decomposition theorem). *Let* $(\mathbf{X}, \mathcal{M}, \mu)$ *be a signed measure space. Then there exist a positive set* $P \in \mathcal{M}$ *and a negative set* $N \in \mathcal{M}$ *such that* $P \cap N = \emptyset$ *and* $P \sqcup N = \mathbf{X}$*. Moreover, for any pair* (P', N') *of such sets, we have that* $P \Delta P' = N \Delta N'$ *is null for* μ*.*

Proof. Without loss of generality, we may assume that μ does not take the value $-\infty$ for we can replace μ by $-\mu$ otherwise. Set

$$r = \inf\{\mu(E) : E \text{ is a negative set for } \mu\}$$

Note that the collection of E's on the right-hand side is clearly non-empty as it contains $\emptyset$. By the definition of infimum, we can choose a sequence E_n of negative sets such that $\lim_{n\to\infty} \mu(E_n) = r$.

Set $N = \bigcup_{n\in\mathbb{N}} E_n$. By Lemma 7, N is a negative set. Moreover, by definition,

$$r \leq \mu(N) \leq \mu(E_n \sqcup (N - E_n)) = \mu(E_n) + \mu(N - E_n) \leq \mu(E_n)$$

for all $n \in \mathbb{N}$ and hence, by taking the limit as $n \to \infty$, we see that $\mu(N) = r$. Set $P = \mathbf{X} - N$. We shall next show that P is a positive set. Assume towards a contradiction that P is not a positive set. Then there exists a measurable set $A \subseteq P$ such that $-\infty < \mu(A) < 0$. By Lemma 8, there exists a negative set $B \subseteq A$ such that $-\infty < \mu(B) \leq \mu(A) < 0$. But then, $N \sqcup B$ is a negative set by Lemma 7 and moreover,

$$\mu(N \sqcup B) = \mu(N) + \mu(B) < \mu(N) = r = \inf\{\mu(E) : E \text{ is a negative set for } \mu\}$$

which is a contradiction. Thus P is a positive set.

Finally, we prove the uniqueness of such decomposition. Let P' be a positive set and N' be a negative set such that $P' \cap N' = \emptyset$ and $P' \sqcup N' = \mathbf{X}$. Then $P \cap N'$ and $P' \cap N$ are both positive and negative sets and so, they are null sets for μ. It follows that $P\Delta P' (= N \Delta N') \subseteq (P \cap N') \cup (P' \cap N)$ is a null set for μ. □

Given a signed measure space $(\mathbf{X}, \mathcal{M}, \mu)$, any pair (P, N) as in Theorem 39 is called a *Hahn decomposition* of the signed measure space $(\mathbf{X}, \mathcal{M}, \mu)$. Although Hahn decompositions are not unique, as stated in Theorem 39, they are unique up to null sets for μ. For this reason, we may say *the* Hahn decomposition to mean any Hahn decomposition.

Next we shall prove that we can decompose a signed measure into two positive measures using the Hahn decomposition of a signed measure space. In order to state this result in a more compact way, we now introduce the notion of mutual singularity.

Given a measurable space $(\mathbf{X}, \mathcal{M})$ and two signed measures ν, η on it, we say that ν and η are *mutually singular* if there exist sets $E, F \in \mathcal{M}$ with $E \cap F = \emptyset$

and $E \sqcup F = \mathbf{X}$ such that E is null for ν and F is null for η. We are now ready to prove the second main result of this section.

Theorem 40 (Jordan decomposition theorem)**.** *Let* $(\mathbf{X}, \mathcal{M}, \mu)$ *be a signed measure space. Then there exist unique positive measures* μ^+ *and* μ^- *such that* $\mu = \mu^+ - \mu^-$ *and* $\mu^+ \perp \mu^-$*.*

Proof. We first prove the existence of such measures. Let (P, N) be the Hahn decomposition of $(\mathbf{X}, \mathcal{M}, \mu)$ which exists by Theorem 39. Define $\mu^+ : \mathcal{M} \to [0, \infty]$ by

$$\mu^+(E) = \mu(E \cap P)$$

for all $E \in \mathcal{M}$. Similarly, define $\mu^- : \mathcal{M} \to [0, \infty]$ by

$$\mu^-(E) = -\mu(E \cap N)$$

for all $E \in \mathcal{M}$. It is straightforward to check that μ^+ and μ^- are (positive) measures. Moreover, for any $E \in \mathcal{M}$, we have that

$$\mu(E) = \mu((E \cap P) \sqcup (E \cap N)) = \mu(E \cap P) + \mu(E \cap N) = \mu^+(E) - \mu^-(E)$$

It is also trivial to see that N is null for μ^+ and P is null for μ^-. Thus we have that $\mu = \mu^+ - \mu^-$ and $\mu^+ \perp \mu^-$.

It remains to prove the uniqueness of these measures. Suppose that $\mu = \nu - \eta$ for some positive measures ν and η with $\nu \perp \eta$. Then, by mutual singularity, we have that $\mathbf{X} = A \sqcup B$ for some disjoint sets $A, B \in \mathcal{M}$ such that A is null for η and B is null for ν. On the other hand, since ν and η are positive measures, A is a positive set and B is a negative set for μ. Consequently, (A, B) is a Hahn decomposition of $(\mathbf{X}, \mathcal{M}, \mu)$. By the uniqueness of the Hahn decomposition, we obtain that $P \Delta A = N \Delta B$ is null for μ. It now follows that

$$\begin{aligned}
\mu^+(E) = \mu(E \cap P) &= \mu((E \cap (P \cap A)) \sqcup (E \cap (P - A))) \\
&= \mu(E \cap (P \cap A)) + \mu(E \cap (P - A)) \\
&= \mu(E \cap (P \cap A)) \\
&= \mu(E \cap (P \cap A)) + \mu(E \cap (A - P)) \\
&= \mu((E \cap (P \cap A)) \sqcup (E \cap (A - P))) \\
&= \mu(E \cap A) = \nu(E \cap A) - \eta(E \cap A) \\
&= \nu(E \cap A) \\
&= \nu(E \cap A) + \nu(E \cap B) = \nu((E \cap A) \sqcup (E \cap B)) = \nu(E)
\end{aligned}$$

for all $E \in \mathcal{M}$. By a similar argument, one can show that $\mu^-(E) = \eta(E)$ for all $E \in \mathcal{M}$. Hence such μ^+ and μ^- are unique. $\square$

The pair (μ^+, μ^-) is called the *Jordan decomposition* of μ. The positive measures μ^+ and μ^- are called the *positive variation* and the *negative variation* of μ respectively. The total variation of μ is the (positive) measure

$$|\mu| = \mu^+ + \mu^-$$

The reader is expected to verify the following properties of the total variation

- For all $E \in \mathcal{M}$, E is $|\mu|$-null if and only if E is null for μ.
- $\mu \perp \nu$ if and only if $|\mu| \perp \nu$.

Let $(\mathbf{X}, \mathcal{M}, \mu)$ be a signed measure space and let (P, N) be its Hahn decomposition. Let (μ^+, μ^-) be the Jordan decomposition of μ. Then we have that

$$\begin{aligned} \mu(E) = \mu^+(E) - \mu^-(E) &= \mu^+(E \cap P) - \mu^-(E \cap N) \\ &= |\mu|(E \cap P) - |\mu|(E \cap N) \\ &= \int_{\mathbf{X}} \chi_{E \cap P} \ d|\mu| - \int_{\mathbf{X}} \chi_{E \cap N} \ d|\mu| \\ &= \int_{\mathbf{X}} \chi_E \cdot (\chi_P - \chi_N) \ d|\mu| \\ &= \int_E (\chi_P - \chi_N) \ d|\mu| \end{aligned}$$

for all $E \in \mathcal{M}$. In other words, as we have mentioned at the beginning of this subsection, every signed measure can be written as a difference of two positive measures and as an integral of a measurable function with respect to a positive measure.

One can build a theory of integration with respect to a signed measure simply by considering integrals with respect its positive and negative variations. However, we shall not proceed in that direction due to time limitations.

5.2. **Radon-Nikodym derivatives.** Let $(\mathbf{X}, \mathcal{M}, \nu)$ be a measure space. Recall that, given a measurable function $f : \mathbf{X} \to [0, \infty]$, the map $\mu : \mathcal{M} \to [0, \infty]$ defined by $\mu(E) = \int_E f \ d\nu$ is a measure by the Monotone Convergence Theorem. It turns out that, under certain hypotheses, one can reverse this procedure and extract such a measurable function f given μ and ν. In order to state the necessary hypothesis, we shall now introduce the notion of absolute continuity of measures with respect to each other.

Let μ and ν be (positive) measures on a measurable space $(\mathbf{X}, \mathcal{M})$. We say that μ is absolutely continuous with respect to ν if

$$\text{for all } E \in \mathcal{M}, \ \mu(E) = 0 \text{ whenever } \nu(E) = 0.$$

In this case, we write $\mu \ll \nu$. Although one can extend this definition to signed measures[1], we shall not need this more general definition for our purposes. What are some examples of measures that are absolutely continuous with respect to others?

Let ν be a measure on a measurable space $(\mathbf{X}, \mathcal{M})$ and let $f : \mathbf{X} \to [0, \infty]$ be a measurable map. Consider the measure $\mu : \mathcal{M} \to [0, \infty]$ given by $\mu(E) = \int_E f \, d\nu$. For any $E \in \mathcal{M}$ with $\nu(E) = 0$, we have that

$$\mu(E) = \int_E f \, d\nu = \int_{\mathbf{X}} \chi_E f \, d\nu = \int_{\mathbf{X}} 0 \, d\nu = 0$$

Thus $\mu \ll \nu$. We shall soon see that all σ-finite examples of measures that are absolutely continuous with respect to others have to be of this form.

Before we prove this, let use give another characterization of absolute continuity for finite measures in terms of an ϵ-δ statement.

Proposition 14. *Let $(\mathbf{X}, \mathcal{M})$ be a measurable space and μ and ν be finite (positive) measures on $(\mathbf{X}, \mathcal{M})$. Then $\mu \ll \nu$ if and only if for every $\epsilon \in \mathbb{R}^+$ there exists $\delta \in \mathbb{R}^+$ such that for every $E \in \mathcal{M}$ we have if $\nu(E) < \delta$, then $\mu(E) < \epsilon$.*

Proof. Assume that the ϵ-δ statement holds. Let $E \in \mathcal{M}$ be such that $\nu(E) = 0$. Let $\epsilon \in \mathbb{R}^+$ be arbitrary. Then there exists $\delta \in \mathbb{R}^+$ such that if $\nu(E) < \delta$, then $\mu(E) < \epsilon$. On the other hand, we have $\nu(E) = 0 < \delta$ and so, $\mu(E) < \epsilon$. We have shown that $\mu(E) < \epsilon$ for any $\epsilon \in \mathbb{R}^+$ and hence, $\mu(E) = 0$. This means that $\mu \ll \nu$.

Now assume that the ϵ-δ statement fails. That is, there exists $\epsilon \in \mathbb{R}^+$ such that for all $\delta \in \mathbb{R}^+$ there exists $E \in \mathcal{M}$ with $\nu(E) < \delta$ and $\mu(E) \geq \epsilon$. Fix such an $\epsilon \in \mathbb{R}^+$. For each $k \in \mathbb{N}^+$, we can choose $E_k \in \mathcal{M}$ such that $\nu(E_k) < 2^{-k}$ and $\mu(E_k) \geq \epsilon$. Set $F_n = \bigcup_{k=n}^{\infty} E_k$ for each $n \in \mathbb{N}^+$. Then, for each $n \in \mathbb{N}^+$, we have that

$$\nu(F_n) \leq \sum_{k=n}^{\infty} \nu\left(E_k\right) \leq \sum_{k=n}^{\infty} 2^{-k} = 2^{1-n}$$

Set $F = \bigcap_{n \in \mathbb{N}^+} F_n$. Then clearly $F \in \mathcal{M}$ and by Theorem 5.d, as ν is a finite measure, we have that

$$\nu(F) = \lim_{n \to \infty} \nu(F_n) \leq \lim_{n \to \infty} 2^{1-n} = 0$$

and that

$$\mu(F) = \lim_{n \to \infty} \mu(F_n) \geq \lim_{n \to \infty} \mu(E_n) \geq \epsilon$$

Therefore, $\nu(F) = 0$ and $\mu(F) \neq 0$. This means that μ is not absolutely continuous with respect to ν. $\square$

Next will be proven the main theorem of this subsection, namely, the Radon-Nikodym theorem.

[1]Given two signed measures μ and ν, we say that μ is absolutely continuous with respect to ν if $|\mu| \ll |\nu|$.

Theorem 41 (The Radon-Nikodym theorem)**.** *Let $(\mathbf{X},\mathcal{M})$ be a measurable space. Let μ and ν be σ-finite measures on $(\mathbf{X},\mathcal{M})$ such that $\mu \ll \nu$. Then there exists a measurable map $g : \mathbf{X} \to [0,\infty)$ such that $\mu(E) = \int_E g \ d\nu$ for all $E \in \mathcal{M}$. Moreover, this map g is unique up to ν-almost everywhere equality.*

Proof. We shall first prove the existence of such a map g in the case that μ and ν are finite measures. Consider the set

$$\mathcal{F} = \left\{ f \in L^+(\mathbf{X},\mathcal{M},\nu) : \int_E f \ d\nu \leq \mu(E) \text{ for all } E \in \mathcal{M} \right\}$$

Observe that

- $\mathcal{F}$ is non-empty as the zero map $\mathbf{0}$ is in $\mathbb{F}$.
- $\mathcal{F}$ is closed under taking maximums. To see this, let $f, \hat{f} \in \mathcal{F}$ and consider $\max\{f,\hat{f}\} \in L^+(\mathbf{X},\mathcal{M},\nu)$. Let $E \in \mathcal{M}$. Set $A = \{x \in E : f(x) \leq \hat{f}(x)\}$ and $B = E - A$. Then clearly $A, B \in \mathcal{M}$ and so, as $f, \hat{f} \in \mathcal{F}$, we have that

$$\int_E \max\{f,\hat{f}\} \ d\nu = \int_A \hat{f} \ d\nu + \int_B f \ d\nu \leq \mu(A) + \mu(B) = \mu(A \sqcup B) = \mu(E)$$

  Therefore $\max\{f,\hat{f}\} \in \mathcal{F}$.

By the definition of $\mathcal{F}$, there exists a sequence $(f_n)_{n\in\mathbb{N}}$ of functions in $\mathcal{F}$ such that

$$\lim_{n\to\infty} \int_{\mathbf{X}} f_n \ d\nu = \sup_{f\in\mathcal{F}} \left\{ \int_{\mathbf{X}} f \ d\nu \right\}$$

Since $\mathcal{F}$ is closed under taking maximums, by setting $g_0 = f_0$ and $g_n = \max\{f_{n+1}, g_n\}$ for all $n \in \mathbb{N}$, we obtain a *monotone* sequence $(g_n)_{n\in\mathbb{N}}$ of functions in $\mathcal{F}$ such that

$$\lim_{n\to\infty} \int_{\mathbf{X}} g_n \ d\nu = \sup_{f\in\mathcal{F}} \left\{ \int_{\mathbf{X}} f \ d\nu \right\}$$

Let $g : \mathbf{X} \to [0,\infty]$ be defined by $g(x) = \lim_{n\to\infty} g_n(x)$ for all $x \in \mathbf{X}$. By the Monotone Convergence Theorem, we have that

$$\int_E g \ d\nu = \lim_{n\to\infty} \int_E g_n \ d\nu \leq \mu(E)$$

for all $E \in \mathcal{M}$. Therefore $g \in \mathcal{F}$. Observe moreover that, by the Monotone Convergence Theorem, we also obtain

$$\int_{\mathbf{X}} g \ d\nu = \lim_{n\to\infty} \int_{\mathbf{X}} g_n \ d\nu = \sup_{f\in\mathcal{F}} \left\{ \int_{\mathbf{X}} f \ d\nu \right\}$$

We shall next prove that we indeed have

$$\int_E g \ d\nu = \mu(E)$$

for all $E \in \mathcal{M}$. We have already proven that $g \in \mathcal{F}$ and so, the left-hand side is less than or equal to the right-hand side. To prove the other inequality, consider the (positive) measure given by $\eta(E) = \mu(E) - \int_E g \ d\nu$ for all $E \in \mathcal{M}$. We wish to show that $\eta(E) = 0$ for all $E \in \mathcal{M}$. Assume towards a contradiction that this is not the case. Then there exists some set $\hat{E} \in \mathcal{M}$ with $\eta(\hat{E}) > 0$ and hence, $\eta(\mathbf{X}) > 0$. As ν is a finite measure, we can find $\epsilon \in \mathbb{R}^+$ such that

$$\eta(\mathbf{X}) > \epsilon\nu(\mathbf{X})$$

Let (P, N) be the Hahn decomposition of the signed measure $\eta - \epsilon\nu$. Let $E \in \mathcal{M}$. Then we have $(\eta - \epsilon\nu)(E \cap P) \geq 0$ and hence,

$$\begin{aligned}\mu(E) = \eta(E) + \int_E g \ d\nu &\geq \eta(E \cap P) + \int_E g \ d\nu \\ &\geq \epsilon\nu(E \cap P) + \int_E g \ d\nu \\ &\geq \int_E \epsilon\chi_P \ d\nu + \int_E g \ d\nu \\ &\geq \int_E g + \epsilon\chi_P \ d\nu\end{aligned}$$

It follows that $g + \epsilon\chi_P \in \mathcal{F}$.

We shall next show that $\nu(P) > 0$. Assume to the contrary that $\nu(P) = 0$. Then, by the assumption that $\mu \ll \nu$, we get $\mu(P) = 0$. This means that

$$0 \leq (\eta - \epsilon\nu)(P) = \eta(P) = \mu(P) - \int_P g \ d\nu = -\int_P g \ d\nu \leq 0$$

which subsequently implies

$$(\eta - \epsilon\nu)(\mathbf{X}) = (\eta - \epsilon\nu)(P) + (\eta - \epsilon\nu)(N) = (\eta - \epsilon\nu)(N) \leq 0$$

This contradicts the choice of ϵ. Therefore $\nu(P) > 0$. Recall that we showed earlier that $\int_{\mathbf{X}} g \ d\nu \leq \mu(\mathbf{X}) < +\infty$. It now follows from $\nu(P) > 0$ that

$$\sup_{f \in \mathcal{F}} \left\{ \int_{\mathbf{X}} f \ d\nu \right\} = \int_{\mathbf{X}} g \ d\nu < \int_{\mathbf{X}} g + \epsilon\chi_P \ d\nu$$

But this is a contradiction as $g + \epsilon\chi_P \in \mathcal{F}$. This completes the proof that $\eta(E) = 0$ for all $E \in \mathcal{M}$. Having proven that

$$\int_E g \ d\nu = \mu(E)$$

for all $E \in \mathcal{M}$, we see that g cannot take the value $+\infty$ on a set having positive measure with respect to ν because $\mu(\mathbf{X}) < +\infty$. Consequently, we can modify the function g on a ν-null set so that its values are contained in $[0, \infty)$ and still satisfies the above equality. This finishes the proof of the existence of such a map g for finite measures.

Now suppose that μ and ν are σ-finite measures. Then we can find a sequence $(X_n)_{n \in \mathbb{N}}$ of disjoint sets in $\mathcal{M}$ such that $\mathbf{X} = \bigsqcup_{n \in \mathbb{N}} X_n$ and $\mu(X_n), \nu(X_n) < \infty$ for all $n \in \mathbb{N}$. Having proven the result for finite measures, in particular for the restrictions μ_n and ν_n of the measures μ and ν to the measurable spaces

$$(X_n, \mathcal{M} \restriction X_n)$$

we can find measurable maps $g_n : X_n \to [0,\infty)$ such that $\mu(E) = \int_E g_n \, d\nu$ for all $E \in \mathcal{M} \upharpoonright X_n$. Consider the measurable map $g : \mathbf{X} \to [0,\infty)$ given by $g(x) = g_n(x)$ if $x \in X_n$. Then, for all $E \in \mathcal{M}$, by Proposition 9, we have that

$$\begin{aligned}
\mu(E) = \mu\left(\bigsqcup_{n\in\mathbb{N}} (X_n \cap E)\right) &= \sum_{n\in\mathbb{N}} \mu(X_n \cap E) \\
&= \sum_{n\in\mathbb{N}} \int_{X_n \cap E} g_n \, d\nu_n \\
&= \sum_{n\in\mathbb{N}} \int_{\mathbf{X}} g \cdot \chi_{X_n \cap E} \, d\nu \\
&= \int_{\mathbf{X}} \sum_{n\in\mathbb{N}} g \cdot \chi_{X_n \cap E} \, d\nu \\
&= \int_{\mathbf{X}} g \cdot \chi_E \, d\nu = \int_E g \, d\nu
\end{aligned}$$

This finishes the proof of the existence of such map g. We now prove the uniqueness up to ν-almost everywhere equality. Let $g, h : \mathbf{X} \to [0,\infty)$ be measurable maps such that

$$\mu(E) = \int_E g \, d\nu = \int_E h \, d\nu$$

for all $E \in \mathcal{M}$. Then, by Lemma 5, we have that $g(x) = h(x)$ holds ν-almost everywhere.[2] $\square$

The map g in Theorem 41, which is unique up to ν-null sets, is called the *Radon-Nikodym derivative of μ with respect to ν* and is shown by

$$g = \frac{d\mu}{d\nu}$$

The motivation behind the terminology and notation is that the Radon-Nikodym derivative is supposed to describe the "rate of change of the density of μ with respect to ν." As the notation suggests, the Radon-Nikodym derivative satisfies many properties of the usual notion of derivative. For example, it is linear and satisfies the chain rule[3], that is, if $\mu \ll \nu$ and $\nu \ll \eta$, then

$$\frac{d\mu}{d\eta} = \frac{d\mu}{d\nu} \cdot \frac{d\nu}{d\eta}$$

Let us now consider some basic examples.

[2]Note that we do not really need to integrability assumption in the proof of the relevant part of Lemma 5.

[3]We shall not prove this fact here for it follows from a problem in Homework III.

- Let $F : \mathbb{R} \to \mathbb{R}$ be an increasing differentiable function and let μ_F be the corresponding Borel measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Suppose that $\mu_F \ll \mathbf{m}$. Then
$$\frac{d\mu_F}{d\mathbf{m}} = F'$$
because we have $\mu_F((a,b]) = F(b) - F(a) = \int_{(a,b]} F' \, d\mathbf{m}$ for all $a < b$ in $\mathbb{R}$.
- Consider the Dirac measure δ_k concentrated at $k \in \mathbb{N}$ and the counting measure μ on the measurable space $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Clearly we have $\delta_k \ll \mu$. Moreover, it is easily seen that
$$\frac{d\delta_0}{d\mu} = \chi_k$$

The Radon-Nikodym theorem can be generalized to signed and complex measures. The reader is referred to [Fol99, Theorem 3.8 and Theorem 3.12] for these more general forms known as *the Lebesgue-Radon-Nikodym theorem*, which essentially tells use how to uniquely "decompose" a σ-finite signed or complex measure into two parts one of which is absolutely continuous with respect to a σ-finite positive measure.

5.3. **Exercises.** Below you shall find some exercises that you can work on regarding the topics in this section. These exercises are *not* to be handed in as homework assignments.

- Exercises 2, 6, 7, 8, 10, 13 from Chapter 3 of [Fol99].
- Exercises 1, 3 from Chapter 4.1 and Exercises 1, 4, 5, 9 from Chapter 4.2 of [Coh93].

6. Coda

This course is intended to serve as a graduate-level introductory course in measure and integration theory. The author hopes that the reader enjoyed the course and benefited as much as possible. Those who wish to study topics in analysis should read the rest of [Fol99], at least, to have an introductory knowledge on functional and Fourier analysis. Those who wish to study topics in abstract measure theory and who seek other books are referred to [Bog07], which is *the* most comprehensive treatment of the subject that the author was able to find.

> "Mathematics, rightly viewed, possesses not only truth, but supreme beauty - a beauty cold and austere, like that of sculpture, without appeal to any part of our weaker nature, without the gorgeous trappings of painting or music, yet sublimely pure, and capable of a stern perfection such as only the greatest art can show. The true spirit of delight, the exaltation, the sense of being more than Man, which is the touchstone of the highest excellence, is to be found in mathematics as surely as poetry."
>
> Bertrand Russell.

References

[Bog07] V. I. Bogachev, *Measure theory. Vol. I, II*, Springer-Verlag, Berlin, 2007. MR 2267655

[Coh93] Donald L. Cohn, *Measure theory*, Birkhäuser Boston, Inc., Boston, MA, 1993, Reprint of the 1980 original. MR 1454121

[Fol99] Gerald B. Folland, *Real analysis*, second ed., Pure and Applied Mathematics (New York), John Wiley & Sons, Inc., New York, 1999, Modern techniques and their applications, A Wiley-Interscience Publication. MR 1681462

[Kec95] Alexander S. Kechris, *Classical descriptive set theory*, Graduate Texts in Mathematics, vol. 156, Springer-Verlag, New York, 1995. MR 1321597

[Lav76] Richard Laver, *On the consistency of Borel's conjecture*, Acta Math. **137** (1976), no. 3-4, 151–169. MR 422027

[Oxt80] John C. Oxtoby, *Measure and category*, second ed., Graduate Texts in Mathematics, vol. 2, Springer-Verlag, New York-Berlin, 1980, A survey of the analogies between topological and measure spaces. MR 584443

[Sie28] W. Sierpiński, *Sur un ensemble non denombrable, dont toute image continue est de mesure nulle*, Fund. Math **11** (1928), no. 1, 302–304.

[SS05] Elias M. Stein and Rami Shakarchi, *Real analysis*, Princeton Lectures in Analysis, vol. 3, Princeton University Press, Princeton, NJ, 2005, Measure theory, integration, and Hilbert spaces. MR 2129625

[Wag93] Stan Wagon, *The Banach-Tarski paradox*, Cambridge University Press, Cambridge, 1993, With a foreword by Jan Mycielski, Corrected reprint of the 1985 original. MR 1251963

Middle East Technical University, Ankara, Turkey,
Email address: burakk@metu.edu.tr