Measure Theory

Lecture notes

Instructor: Arijit Chakrabarty

Last updated: 6 January 2026

1 Motivation for learning Measure Theory

Two important branches of mathematics, namely analysis and probability, are built on measure theory. The following subsections talk about the motivations from these two branches separately.

1.1 Analytic motivation

In analysis, the theory of Riemann integration was found to have a few shortcomings, despite giving a rigorous definition of the integral in terms of Riemann sums which has its intuitive appeal. One of the first shortcomings was it failed to integrate functions like the following. Define $f : [0, 1] \to \mathbb{R}$ by

$$f(x) = \begin{cases} 0, & x \in \mathbb{Q} \cap [0, 1], \\ 1, & \text{else}. \end{cases} \tag{1.1.1}$$

A moment's thought reveals that the integral of the above function on $[0, 1]$ should be 1 because its value is 1 "almost everywhere", since there are only countably many rational numbers.

The second shortcoming, which arose out of technical requirements, is that the space of Riemann integrable functions is not "complete". To make this precise, let $\mathcal{R}$ be the collection of bounded Riemann integrable functions from $[0, 1]$ to $\mathbb{R}$ equipped with the pseudo-metric d defined by

$$d(f, g) = \int_0^1 |f(x) - g(x)|\, dx,\ f, g \in \mathcal{R}\,.$$

A metric space can be constructed out of $(\mathcal{R}, d)$ by identifying $f, g \in \mathcal{R}$ whenever $d(f, g) = 0$. This metric space is not complete, that is, there exist Cauchy sequences under d which are not convergent.

Another shortcoming of the Riemann theory is that the fundamental theorem of calculus, which says the following for a differentiable function $f : [a, b] \to \mathbb{R}$ with derivative f':

$$\int_a^b f'(x)\, dx = f(b) - f(a), \tag{1.1.2}$$

fails whenever f' is not Riemann integrable on $[a, b]$ because then the left hand side above doesn't make sense.

For the reasons mentioned above, the Riemann theory was considered insufficient and there was a necessity to integrate a larger class of functions. In other words, a richer theory of integration was called for. This was done by Henri Lebesgue in his doctoral thesis on "Integral, Length, Area", which appeared in the Annali di Matematica in 1902. Lebesgue's work can be regarded as the starting point of the study of measure theory. Indeed, Lebesgue's theory of measure and integration solved the first two shortcomings of the Riemann theory mentioned above. Further, (1.1.2) holds for a larger class of functions f, when its left hand side is considered in the Lebesgue sense, though not for every differentiable function.

1.2 Probabilistic motivation

Probability theory starts with the classical definition of probability, in which, the probability of an event is defined as the ratio of its cardinality and that of the sample space, which is the set of all possible outcomes. While this definition makes perfect sense as long as the sample space is finite and all the outcomes therein are "equally likely", it fails the moment either the sample space becomes infinite or the outcomes do not remain equally likely. Although the definition can be stretched a bit to accommodate the case when the sample space is countable, it is insufficient to deal with an uncountable sample space. For example, if a (fair) coin is tossed (countably) infinitely many times, then the sample space

$$\Omega = \big\{(\omega_1, \omega_2, \ldots) : \omega_n \in \{H, T\} \text{ for all } n = 1, 2, 3, \ldots\big\}$$

is uncountable. A firm mathematical framework for studying such random experiments was provided by measure theory.

2 Measure

A close look at f as in (1.1.1) tells us that in principle, its integral should be defined as $0\cdot\lambda(\mathbb{Q}\cap[0, 1])+1\cdot\lambda(\mathbb{Q}^c\cap[0, 1])$, where $\lambda(A)$ is the "length" of a set A. In other words, our starting point should be to generalize the concept of length of an interval to "measure" of all subsets. Unfortunately, this cannot be done for all subsets, which is proved in the following subsection.

2.1 The Vitali construction

For any set X, its power set is denoted throughout by 2^X.

Theorem 2.1.1. *There does not exist a function* $\lambda : 2^{[0,1)} \to [0,1]$ *such that* $\lambda([0,1)) = 1$ *and* λ *is*

1. *countably additive, that is, for disjoint* $A_1, A_2, A_3, \ldots \subset [0,1)$,
$$\lambda\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \lambda(A_n),$$
2. *and translation invariant, that is, if* $A \subset [0,1)$ *and* $x \in \mathbb{R}$ *are such that*
$$A + x := \{a + x : a \in A\} \subset [0,1),$$
then $\lambda(A + x) = \lambda(A)$.

The proof uses the **axiom of choice**, which says that for any collection $\mathcal{S}$ of non-empty sets, there exists a function
$$f : \mathcal{S} \to \bigcup_{A \in \mathcal{S}} A$$
such that $f(A) \in A$ for all $A \in \mathcal{S}$.

Proof of Theorem 2.1.1. Assume for the sake of contradiction such λ exists. For $x, y \in [0,1)$, say $x \sim y$ if $x - y \in \mathbb{Q}$. Then $\sim$ is an equivalence relation. Let $\mathcal{S}$ be the collection of equivalence classes under $\sim$, that is,
$$\mathcal{S} \subset 2^{[0,1)} \setminus \{\emptyset\},$$
$$\bigcup_{A \in \mathcal{S}} A = [0,1), \tag{2.1.1}$$
for $A, B \in \mathcal{S}$, either $A = B$ or $A \cap B = \emptyset$, and
$$\text{for all } x, y \in \mathcal{S},\ x \sim y \iff x, y \in A \text{ for some } A \in \mathcal{S}\,. \tag{2.1.2}$$

The axiom of choice implies there exists a function $f : \mathcal{S} \to [0,1)$ such that
$$f(A) \in A \text{ for all } A \in \mathcal{S}\,.$$
Define
$$V = f(\mathcal{S})\,. \tag{2.1.3}$$
Since this construction is due to Vitali, V will be later referred to as the "Vitali set". The definition of V ensures that
$$\#(V \cap A) = 1 \text{ for all } A \in \mathcal{S}\,. \tag{2.1.4}$$

For $x, y \in [0,1)$, define

$$x \oplus y = \begin{cases} x+y, & \text{if } x+y<1, \\ x+y-1, & \text{if } x+y \geq 1\,. \end{cases}$$

Trivially, $x \oplus y \in [0,1)$ for all $x, y \in [0,1)$. Hence for $x \in [0,1)$,

$$V \oplus x := \{v \oplus x : v \in V\} \subset [0,1)\,.$$

Our first claim is that

$$\lambda(V \oplus x) = \lambda(V),\ x \in [0,1)\,. \tag{2.1.5}$$

To see the above, write

$$V \oplus x = \big((V \cap [0, 1-x)) + x\big) \cup \big((V \cap [1-x, 1)) + x - 1\big),$$

and observe that

$$\begin{aligned} (V \cap [0, 1-x)) + x &\subset [x, 1), \\ (V \cap [1-x, 1)) + x - 1 &\subset [0, x)\,. \end{aligned}$$

Countable additivity of λ implies

$$\begin{aligned} \lambda(V \oplus x) &= \lambda\big((V \cap [0, 1-x)) + x\big) + \lambda\big((V \cap [1-x, 1)) + x - 1\big) \\ &= \lambda\big((V \cap [0, 1-x))\big) + \lambda\big((V \cap [1-x, 1))\big) \\ &= \lambda(V), \end{aligned}$$

the second line being implied by translation invariance of λ and the last line again following from countable additivity; (2.1.5) thus follows.

We shall next show that

$$\bigcup_{r \in \mathbb{Q} \cap [0,1)} (V \oplus r) = [0,1)\,. \tag{2.1.6}$$

Fix $x \in [0,1)$. It follows from (2.1.1) that there exists $A \in \mathcal{S}$ such that $x \in A$. A consequence of (2.1.4) is that $A \cap V$ is a singleton set, say

$$A \cap V = \{v\}\,.$$

Thus x and v are in the same equivalence class which is A, and hence (2.1.2) implies $x \sim v$. In other words, $x - v \in \mathbb{Q}$. Define

$$r = \begin{cases} x - v, & \text{if } v \leq x, \\ x - v + 1, & \text{if } v > x\,. \end{cases}$$

It is easy to check that $r \in \mathbb{Q} \cap [0,1)$ and $x = v \oplus r$. Since $v \in V$, we get $x \in V \oplus r$. Thus

$$\bigcup_{r \in \mathbb{Q} \cap [0,1)} (V \oplus r) \supset [0,1)\,.$$

The reverse inclusion being trivial, (2.1.6) follows.

We shall next prove that the left hand side of (2.1.6) is a disjoint union, that is,

$$(V \oplus r) \cap (V \oplus s) = \emptyset \text{ for distinct } r, s \in \mathbb{Q} \cap [0,1)\,. \tag{2.1.7}$$

Fix distinct $r, s \in \mathbb{Q} \cap [0,1)$ and if possible, let $x \in (V \oplus r) \cap (V \oplus s)$. Thus there exist $y, z \in V$ such that

$$y \oplus r = x = z \oplus s\,.$$

Since $y \oplus r$ equals either $y + r$ or $y + r - 1$ and likewise for $z \oplus s$, one of the following three must hold:

$$y + r = z + s, \tag{2.1.8}$$

$$y + r = z + s - 1, \tag{2.1.9}$$

$$\text{or, } y + r - 1 = z + s\,. \tag{2.1.10}$$

In all the above cases, $y - z \in \mathbb{Q}$, that is, $y \sim z$. By (2.1.2), there exists $A \in \mathcal{S}$ such that $y, z \in A$. Hence $y, z \in A \cap V$; (2.1.4) implies $y = z$. Since $r, s \in [0,1)$, neither (2.1.9) nor (2.1.10) can hold because $|r - s| < 1$. Therefore, (2.1.8) holds, which shows $r = s$ and thus leads to a contradiction. This proves (2.1.7).

Countable additivity of λ in conjunction with (2.1.6) and (2.1.7) shows

$$\begin{aligned} \lambda\big([0,1)\big) &= \sum_{r \in \mathbb{Q} \cap [0,1)} \lambda(V \oplus r) \\ &= \sum_{r \in \mathbb{Q} \cap [0,1)} \lambda(V), \end{aligned}$$

(2.1.5) implying the second line. Obviously,

$$\sum_{r \in \mathbb{Q} \cap [0,1)} \lambda(V) = \begin{cases} 0, & \text{if } \lambda(V) = 0, \\ \infty, & \text{if } \lambda(V) > 0\,. \end{cases}$$

This contradicts $\lambda([0,1)) = 1$ and thus completes the proof. ☐

2.2 Definition of measure and its properties

Theorem 2.1.1 tells us that the domain of a "measure" on a set X has to be much smaller than 2^X. This calls for the following definition.

Definition 2.2.1. *Given a non-empty set* Ω, $\mathcal{A} \subset 2^\Omega$ *is a* σ-field on Ω *if* $\Omega \in \mathcal{A}$, $A \in \mathcal{A}$ *implies* $A^c \in \mathcal{A}$, *and* $A_1, A_2, \ldots \in \mathcal{A}$ *implies* $\bigcup_{n=1}^\infty A_n \in \mathcal{A}$. *If* $\mathcal{A}$ *is a* σ*-field on* Ω, *then* $(\Omega, \mathcal{A})$ *is a* measurable space.

Now we are in a position to define measure. Before that let us fix the convention for adding and subtracting ∞. Define

$$x + \infty = \infty + x = \infty \text{ for all } x \in (-\infty, \infty]$$

and

$$x - \infty = -\infty + x = -\infty \text{ for all } x \in [-\infty, \infty)\,.$$

Neither $\infty - \infty$ nor $-\infty + \infty$ is defined. For $a_n \in \overline{\mathbb{R}} = [-\infty, \infty]$, say $a_n \to \infty$ if for all $M \in \mathbb{R}$, there exists N such that

$$a_n > M \text{ for all } n \ge N\,.$$

Similarly, $a_n \to -\infty$ if for all $M \in \mathbb{R}$, there exists N such that

$$a_n < M \text{ for all } n \ge N\,.$$

In view of the above definition, it is easy to check that for $a_1, a_2, \ldots \in [0, \infty]$, there exists $s \in [0, \infty]$ such that

$$\sum_{i=1}^n a_i \to s,\ n \to \infty\,.$$

For such a_n and s, we define

$$\sum_{n=1}^\infty a_n = s\,.$$

Definition 2.2.2. *Given a measurable space* $(\Omega, \mathcal{A})$, *a function* $\mu : \mathcal{A} \to [0, \infty]$ *is a* measure *if* $\mu(\emptyset) = 0$ *and* μ *is countably additive, that is,*

$$\mu(A_1 \cup A_2 \cup \ldots) = \sum_{n=1}^\infty \mu(A_n) \textit{ for all disjoint } A_1, A_2, A_3, \ldots \in \mathcal{A}\,.$$

The tuple $(\Omega, \mathcal{A}, \mu)$ *is a* measure space. *We say* μ *is a*

- probability measure *if* $\mu(\Omega) = 1$ *and in this case* $(\Omega, \mathcal{A}, \mu)$ *is a probability space (a probability measure is usually denoted by* P*),*
- finite measure *if* $\mu(\Omega) < \infty$ *and in this case* $(\Omega, \mathcal{A}, \mu)$ *is a finite measure space,*

- σ-finite measure *if there exist* $A_1, A_2, \ldots \in \mathcal{A}$ *such that*

$$\Omega = \bigcup_{n=1}^{\infty} A_n \text{ and } \mu(A_n) < \infty,\, n \geq 1,$$

  and in this case $(\Omega, \mathcal{A}, \mu)$ *is a* σ*-finite measure space.*

Exercise 2.2.1. *Suppose* $(\Omega, \mathcal{A}, \mu)$ *is a measure space. If* $A \subset B$ *and* $A, B \in \mathcal{A}$*, show that* $\mu(A) \leq \mu(B)$*.*

Exercise 2.2.2. *Suppose* $(\Omega, \mathcal{A}, \mu)$ *is a* σ*-finite measure space. Show that there exist disjoint* $A_1, A_2, \ldots \in \mathcal{A}$ *such that*

$$\Omega = \bigcup_{n=1}^{\infty} A_n \text{ and } \mu(A_n) < \infty,\, n \geq 1\,.$$

Let us give a few examples of measure spaces.

Example 2.2.1. *Let* Ω *be an uncountable set. Define*

$$\mathcal{A} = \{A \subset \mathcal{A} : \text{Either } A \text{ or } A^c \text{ is countable.}\}\,.$$

Clearly, $\mathcal{A}$ *is a* σ*-field, usually called the* countable-cocountable σ*-field. A set is called cocountable if its complement is countable. Fix* $0 \leq \alpha \leq \infty$ *and let*

$$\mu(A) = \begin{cases} \alpha, & \text{if } A \text{ is countable,} \\ 0, & \text{if } A \text{ is cocountable.} \end{cases}$$

Then μ *is a measure on* $(\Omega, \mathcal{A})$*. Furthermore,* μ *is not* σ*-finite if* $\alpha = \infty$*.*

Example 2.2.2. *Suppose* $\Omega \supset \boldsymbol{C} \neq \emptyset$*. Define*

$$\mu(A) = \#(A \cap \boldsymbol{C}),\, A \in 2^{\Omega}\,.$$

Then μ *is a measure on* $(\Omega, 2^{\Omega})$*, called the "counting measure" on* $\boldsymbol{C}$*. If* $\boldsymbol{C}$ *is countable, then* μ *is* σ*-finite.*

The following property of a measure is known as "continuity".

Theorem 2.2.1. *Suppose* $(\Omega, \mathcal{A}, \mu)$ *is a measure space. For* $A_1, A_2, \ldots \in \mathcal{A}$*,*

- *(continuity from below)* $A_n \uparrow A$ *implies* $\mu(A_n) \uparrow \mu(A)$*,*
- *(continuity from above)* $A_n \downarrow A$ *and* $\mu(A_1) < \infty$ *imply* $\mu(A_n) \downarrow \mu(A)$*.*

Proof. Assume first that $A_n \uparrow A$. That $\mu(A_1) \leq \mu(A_2 \leq \ldots$ follows from monotonicity of μ. Denoting $A_0 = \emptyset$, it is immediate that

$$A = \bigcup_{n=1}^{\infty} (A_n \setminus A_{n-1})\,,$$

and that the sets on the right hand side are disjoint. Thus,

$$\begin{aligned}\mu(A) &= \sum_{n=1}^{\infty} \mu\left(A_n \setminus A_{n-1}\right) \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} \mu\left(A_i \setminus A_{i-1}\right) \\ &= \lim_{n\to\infty} \mu(A_n),\end{aligned}$$

the last line following from the observation that

$$A_n = \bigcup_{i=1}^{n} (A_i \setminus A_{i-1})\,.$$

Thus $\mu(A_n) \uparrow \mu(A)$, that is, continuity from below follows.

Now suppose that $A_n \downarrow A$ and $\mu(A_1) < \infty$. Once again, $\mu(A_1) \geq \mu(A_2) \geq \ldots$ follows from monotonicity. Clearly, $A_n^c \uparrow A^c$ and hence $(A_1 \setminus A_n) \uparrow (A_1 \setminus A)$. Continuity of μ from below implies that

$$\mu(A_1 \setminus A_n) \uparrow \mu(A_1 \setminus A)\,.$$

Since $\mu(A_1) < \infty$, the above is the same as

$$\mu(A_1) - \mu(A_n) \uparrow \mu(A_1) - \mu(A)\,.$$

Thus $\mu(A_n) \downarrow \mu(A)$. This completes the proof. □

Exercise 2.2.3. *If $\Omega = \mathbb{N}$ and μ is the counting measure on $\mathbb{N}$, show that*

$$\{n, n+1, \ldots\} \downarrow \emptyset$$

and

$$\mu\left(\{n, n+1, \ldots\}\right) \not\to 0,\ n \to \infty\,.$$

Definition 2.2.3. *For $\Omega \neq \emptyset$ and $\mathcal{G} \subset 2^{\Omega}$, the σ*-field generated by *$\mathcal{G}$, denoted by $\sigma(\mathcal{G})$, is defined as*

$$\sigma(\mathcal{G}) = \bigcap_{\mathcal{A}:\, \mathcal{G} \subset \mathcal{A} \subset 2^{\Omega},\, \mathcal{A} \text{ is a } \sigma\text{-field}} \mathcal{A}\,.$$

In other words, $\sigma(\mathcal{G})$ is the intersection of all σ-fields on Ω which contain $\mathcal{G}$.

Exercise 2.2.4. *If Ω and $\mathcal{G}$ are as above, show that $\sigma(\mathcal{G})$ is the smallest σ-field containing $\mathcal{G}$, that is,*

1. *$\mathcal{G} \subset \sigma(\mathcal{G})$,*

2. *$\sigma(\mathcal{G})$ is a σ-field,*

3. *and if $\mathcal{A}$ is any σ-field on Ω with $\mathcal{A} \supset \mathcal{G}$, then $\sigma(\mathcal{G}) \subset \mathcal{A}$.*

Definition 2.2.4. *The σ-field on $\mathbb{R}$ generated by the collection of all open subsets of $\mathbb{R}$ is called the* Borel σ-field on $\mathbb{R}$ *and is denoted by $\mathcal{B}(\mathbb{R})$. Elements of $\mathcal{B}(\mathbb{R})$ are called* Borel sets.

Exercise 2.2.5. *Show that $A \subset \mathbb{R}$ is a Borel set if A is*

1. *an open set*

2. *a closed set*

3. *a countable set*

4. *or an interval, that is, A is one of (a, b), $(a, b]$, $[a, b)$ or $[a, b]$ for some $-\infty < a \leq b < \infty$.*

2.3 The Carathéodory extension theorem

For constructing useful measures, the usual method is to first define a "countably additive set function" on a "field", and then use an extension theorem to extend it to the generated σ-field. This is done with the help of the Carathéodory extension theorem, stated and proved below. This result is therefore of fundamental importance in measure theory.

Definition 2.3.1. *For $\Omega \neq \emptyset$, $\mathcal{F} \subset 2^\Omega$ is a* field *if $\emptyset \in \mathcal{F}$, $A \in \mathcal{F}$ implies $A^c \in \mathcal{F}$ and $A, B \in \mathcal{F}$ implies $A \cup B \in \mathcal{F}$.*

Exercise 2.3.1. *If $\mathcal{F}$ is a field, show that for $A_1, \ldots, A_n \in \mathcal{F}$,*

$$A_1 \cup \ldots \cup A_n \in \mathcal{F} \text{ and } A_1 \cap \ldots \cap A_n \in \mathcal{F}\,.$$

Definition 2.3.2. *Given a field $\mathcal{F}$ on $\Omega \neq \emptyset$, a function $\mu : \mathcal{F} \to [0, \infty]$ is a* countably additive set function *if $\mu(\emptyset) = 0$ and*

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

whenever $A_1, A_2, \ldots \in \mathcal{F}$ are disjoint and $A_1 \cup A_2 \cup \ldots \in \mathcal{F}$.

Theorem 2.3.1 (Carathéodory extension theorem)**.** *If $\mathcal{F}$ is a field on $\Omega \neq \emptyset$ and μ is a countably additive set function on $\mathcal{F}$, then there exists a measure μ^* on $\big(\Omega, \sigma(\mathcal{F})\big)$ such that*

$$\mu^*(A) = \mu(A) \text{ for all } A \in \mathcal{F}\,.$$

Let μ be a countably additive set function on $\mathcal{F}$. We start with defining μ^* as the "outer measure" of μ as follows:

$$\mu^*(E) = \inf\left\{\sum_{n=1}^\infty \mu(A_n) : E \subset \bigcup_{n=1}^\infty A_n \text{ and } A_1, A_2, \ldots \in \mathcal{F}\right\}, \, E \in 2^\Omega \,. \tag{2.3.1}$$

Lemma 2.3.1. *For all $A \in \mathcal{F}$, $\mu^*(A) = \mu(A)$.*

Proof. Fix $A \in \mathcal{F}$. Letting $A_1 = A$ and $A_2 = A_3 = \ldots = \emptyset$, it is immediate that

$$\mu^*(A) \le \sum_{n=1}^\infty \mu(A_n) = \mu(A)\,.$$

For the reverse inequality, suppose that

$$A \subset \bigcup_{n=1}^\infty A_n \text{ for some } A_1, A_2, \ldots \in \mathcal{F}\,. \tag{2.3.2}$$

Define $B_1 = A \cap A_1$, $B_2 = A \cap A_2 \cap A_1^c$, $B_3 = A \cap A_3 \cap (A_1 \cup A_2)^c$ and in general

$$B_n = A \cap A_n \cap \left(\bigcup_{i=1}^{n-1} A_i\right)^c \text{ for all } n \ge 1\,.$$

Since $\mathcal{F}$ is a field, $B_1, B_2, \ldots \in \mathcal{F}$. Further, (2.3.2) shows

$$B_1 \cup B_2 \cup \ldots = A\,.$$

Obviously, $B_1, B_2, \ldots$ are disjoint. Countable additivity of μ implies that

$$\mu(A) = \sum_{n=1}^\infty \mu(B_n) \le \sum_{n=1}^\infty \mu(A_n),$$

the inequality following from the trivial fact that any countably additive set function is necessarily monotone. Since this holds for any $A_1, A_2, \ldots$ satisfying (2.3.2), we get

$$\mu(A) \le \mu^*(A)\,.$$

This along with the already proven reverse inequality completes the proof. ☐

Lemma 2.3.2. *For all $A_1, A_2, \ldots \in 2^\Omega$,*

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

In other words, μ^* is countably subadditive.

Proof. Let $A_1, A_2, \ldots \in 2^\Omega$. The claim would follow if it can be shown that for all $\varepsilon > 0$,

$$\mu^*\left(\bigcup_{n=1}^\infty A_n\right) \le \varepsilon + \sum_{n=1}^\infty \mu^*(A_n)\,. \tag{2.3.3}$$

Fix $\varepsilon > 0$. By definition of μ^*, there exists $A_{n1}, A_{n2}, \ldots \in \mathcal{F}$ such that

$$A \subset A_{n1} \cup A_{n2} \cup \ldots \text{ and } \sum_{i=1}^\infty \mu(A_{ni}) \le \mu^*(A_n) + 2^{-n}\varepsilon\,.$$

Thus

$$\bigcup_{n=1}^\infty A_n \subset \bigcup_{n=1}^\infty \bigcup_{i=1}^\infty A_{ni}\,.$$

Hence

$$\mu^*\left(\bigcup_{n=1}^\infty A_n\right) \le \sum_{n=1}^\infty \sum_{i=1}^\infty \mu(A_{ni}) \le \sum_{n=1}^\infty \left(\mu^*(A_n) + 2^{-n}\varepsilon\right) = \varepsilon + \sum_{n=1}^\infty \mu^*(A_n)\,.$$

This shows (2.3.3) from which the proof follows. ☐

Lemma 2.3.3. *Define*

$$\mathcal{A} = \{A \in 2^\Omega : \mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \text{ for all } E \in 2^\Omega\}\,. \tag{2.3.4}$$

Then $\mathcal{A} \supset \mathcal{F}$.

Proof. Lemma 2.3.2 shows that

$$\mu^*(E) \le \mu^*(E \cap A) + \mu^*(E \cap A^c) \text{ for all } E, A \in 2^\Omega\,.$$

Hence (2.3.4) becomes

$$\mathcal{A} = \{A \in 2^\Omega : \mu^*(E) \ge \mu^*(E \cap A) + \mu^*(E \cap A^c) \text{ for all } E \in 2^\Omega\}\,. \tag{2.3.5}$$

Let $A \in \mathcal{F}$ and $E \in 2^\Omega$. Fix $\varepsilon > 0$ and let $A_1, A_2, \ldots \in \mathcal{F}$ be such that

$$E \subset \bigcup_{n=1}^\infty A_n \text{ and } \mu^*(E) + \varepsilon \ge \sum_{n=1}^\infty \mu(A_n)\,.$$

The definition of μ^* implies

$$\mu^*(E \cap A) \le \sum_{n=1}^\infty \mu(A_n \cap A),$$

and

$$\mu^*(E \cap A^c) \le \sum_{n=1}^\infty \mu(A_n \cap A^c)\,.$$

Combining the two yields

$$\begin{aligned}\mu^*(E\cap A)+\mu^*(E\cap A^c) &\le \sum_{n=1}^\infty \mu(A_n\cap A)+\sum_{n=1}^\infty \mu(A_n\cap A^c)\\ \text{(countable additivity of } \mu) &= \sum_{n=1}^\infty \mu(A_n)\\ &\le \varepsilon+\mu^*(E)\,.\end{aligned}$$

Since ε is arbitrary, it follows that

$$\mu^*(E\cap A)+\mu^*(E\cap A^c)\le\mu^*(E)\,.$$

As this holds for all $E\in 2^\Omega$, (2.3.5) shows $A\in\mathcal{A}$ and hence completes the proof. $\square$

Lemma 2.3.4. *The collection $\mathcal{A}$, as in* (2.3.4)*, is a field and*

$$\mu^*\left(E\cap\left(\bigcup_{n=1}^\infty A_n\right)\right)=\sum_{n=1}^\infty \mu^*(E\cap A_n) \textit{ for disjoint } A_1,A_2,\ldots\in\mathcal{A},\ E\in 2^\Omega\,.$$

Proof. The definition (2.3.4) shows $\Omega\in\mathcal{A}$ and that $\mathcal{A}$ is closed under complements. Suppose $A,B\in\mathcal{A}$ and $E\in 2^\Omega$. Define

$$F=E\cap(A\cup B)\,.$$

Since $A\in\mathcal{A}$,

$$\mu^*(F)=\mu^*(F\cap A)+\mu^*(F\cap A^c)=\mu^*(F\cap A)+\mu^*(E\cap B\cap A^c)\,.$$

Since $B\in\mathcal{A}$ and $F\cap A\in 2^\Omega$,

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

which implies

$$\mu^*\left(E\cap(A\cup B)\right)=\mu^*(E\cap A\cap B)+\mu^*(E\cap A\cap B^c)+\mu^*(E\cap B\cap A^c)\,. \tag{2.3.6}$$

Once again, $A\in\mathcal{A}$ implies

$$\begin{aligned}\mu^*(E) &= \mu^*(E\cap A)+\mu^*(E\cap A^c)\\ \text{(as } B\in\mathcal{A}) &= \mu^*(E\cap A\cap B)+\mu^*(E\cap A\cap B^c)+\mu^*(E\cap A^c\cap B)\\ &\quad+\mu^*(E\cap A^c\cap B^c)\\ \text{(by (2.3.6))} &= \mu^*\left(E\cap(A\cup B)\right)+\mu^*(E\cap A^c\cap B^c)\\ &= \mu^*\left(E\cap(A\cup B)\right)+\mu^*\left(E\cap(A\cup B)^c\right)\,.\end{aligned}$$

Since this holds for all E, $A \cup B \in \mathcal{A}$, showing that $\mathcal{A}$ is a field.

Fix $E \in 2^\Omega$ and disjoint $A_1, A_2, \ldots \in \mathcal{A}$. Let

$$F = E \cap (A_1 \cup A_2 \cup \ldots)\,.$$

Then

$$\begin{aligned}\mu^*(F) &= \mu^*(F \cap A_1) + \mu^*(F \cap A_1^c)\\ &= \mu^*(E \cap A_1) + \mu^*\left(E \cap (A_2 \cup A_3 \cup \ldots)\right)\,.\end{aligned}$$

Proceeding inductively, it can be shown that for all $n = 1, 2, \ldots$,

$$\mu^*(F) = \sum_{i=1}^n \mu^*(E \cap A_i) + \mu^*\left(E \cap (A_{n+1} \cup \ldots)\right) \geq \sum_{i=1}^n \mu^*(E \cap A_i)\,.$$

Since this is true for all n, we get

$$\mu^*\left(E \cap (A_1 \cup A_2 \cup \ldots)\right) \geq \sum_{i=1}^\infty \mu^*(E \cap A_i)\,.$$

Lemma 2.3.2 yields the reverse inequality and thus completes the proof. ☐

Lemma 2.3.5. *The collection $\mathcal{A}$ is a σ-field and μ^* is a measure on* $(\Omega, \mathcal{A})$.

Proof. Since $\mathcal{A}$ has already been shown to be a field in Lemma 2.3.4, to show $\mathcal{A}$ is a σ-field it suffices to prove it is closed under disjoint countable union. Let $A_1, A_2, \ldots \in \mathcal{A}$ be disjoint. Let

$$A = A_1 \cup A_2 \cup \ldots \text{ and } F_n = A_1 \cup \ldots \cup A_n,\ n \geq 1\,.$$

Since $F_n \in \mathcal{A}$ by Lemma 2.3.4, for any $E \in 2^\Omega$,

$$\begin{aligned}\mu^*(E) &= \mu^*(E \cap F_n) + \mu^*(E \cap F_n^c)\\ \text{(by Lemma 2.3.4)} &= \sum_{i=1}^n \mu^*(E \cap A_i) + \mu^*(E \cap F_n^c)\\ &\geq \sum_{i=1}^n \mu^*(E \cap A_i) + \mu^*(E \cap A^c),\end{aligned}$$

the last line following from the observation that $F_n^c \subset A^c$ and the fact that μ^* is monotone, that is,

$$\mu^*(A) \leq \mu^*(B) \text{ if } A \subset B \subset \Omega,$$

which follows from the very definition of μ^*. Let $n \to \infty$ to get

$$\mu^*(E) \geq \sum_{i=1}^\infty \mu^*(E \cap A_i) + \mu^*(E \cap A^c) = \mu^*(E \cap A) + \mu^*(E \cap A^c),$$

the second equality following once again from Lemma 2.3.4. Thus $A \in \mathcal{A}$ by (2.3.5). Hence $\mathcal{A}$ is a σ-field. Taking $E = \Omega$ in Lemma 2.3.4 shows μ^* is a measure. This completes the proof. ☐

Proof of Theorem 2.3.1. Lemmas 2.3.1, 2.3.3 and 2.3.5 complete the proof. □

2.4 Uniqueness of the extension

After Theorem 2.3.1, the most pertinent question which arises subsequently is whether the extension is unique. The following theorem answers this.

Theorem 2.4.1 (Uniqueness)**.** *If μ and ν are finite measures on* $(\Omega, \sigma(\mathcal{F}))$*, where $\mathcal{F}$ is a field, and*

$$\mu(A) = \nu(A) \text{ for all } A \in \mathcal{F},$$

then μ, ν agree on $\sigma(\mathcal{F})$.

The proof uses the following definition and the "monotone class theorem" which will be proved a moment later.

Definition 2.4.1. *For $\Omega \neq \emptyset$, $\mathcal{M} \subset 2^\Omega$ is a* monotone class *if*

$$A_n \in \mathcal{M} \text{ and } A_n \uparrow A \text{ imply } A \in \mathcal{M},$$

and

$$A_n \in \mathcal{M} \text{ and } A_n \downarrow A \text{ imply } A \in \mathcal{M}\,.$$

Theorem 2.4.2 (Monotone class theorem)**.** *If $\mathcal{F}$ is a field, $\mathcal{F} \subset \mathcal{M}$ and $\mathcal{M}$ is a monotone class, then*

$$\sigma(\mathcal{F}) \subset \mathcal{M}\,.$$

Although 'MCT' is an abbreviation for the above theorem, the same is being saved for the monotone convergence theorem to be stated and proved later.

Proof of Theorem 2.4.1. Define

$$\mathcal{G} = \{A \in \sigma(\mathcal{F}) : \mu(A) = \nu(A)\}\,.$$

The hypothesis implies $\mathcal{F} \subset \mathcal{G}$. Finiteness of $\mu(\Omega)$ and $\nu(\Omega)$ in conjunction with Theorem 2.2.1 implies that

$$\mu(A_n) \to \mu(A) \text{ and } \nu(A_n) \to \nu(A)$$

if either $A_n \uparrow A$ or $A_n \downarrow A$. Thus $A \in \mathcal{G}$ if $A_n \in \mathcal{G}$ and either $A_n \uparrow A$ or $A_n \downarrow A$. In other words, $\mathcal{G}$ is a monotone class. Theorem 2.4.2 shows $\mathcal{G} \supset \sigma(\mathcal{F})$, from which the proof follows. □

Remark 2.4.1. *The technique employed in the proof of Theorem 2.4.1 is a variant of the so-called "good set principle" and is ubiquitous in measure theory, which is the following. For showing that a property holds for every set in a σ-field generated by a collection $\mathcal{H}$, we first define a set to be "good" if it has the said property and then show the following.*

- *The collection of all good sets, say $\mathcal{G}$, is a σ-field*
- *and that $\mathcal{H} \subset \mathcal{G}$.*

The above would ensure from the definition of a σ-field generated by $\mathcal{H}$ that

$$\sigma(\mathcal{H}) \subset \mathcal{G}$$

which is tautologically the same as that the desired property holds for every set in $\sigma(\mathcal{H})$. The tools needed for showing $\mathcal{G}$ to be a σ-field vary with context. In the proof of Theorem 2.4.1, for example, the monotone class theorem was used to essentially show that $\mathcal{G}$ is a σ-field. In other situations, $\mathcal{G}$ is shown to be a σ-field from first principles.

The following stronger version of Theorem 2.4.1 can be proven along similar lines, whose proof is left as an exercise.

Theorem 2.4.3. *Suppose $(\Omega, \mathcal{A})$ is a measurable space on which, μ_1, μ_2 are measures. Assume $\mathcal{F} \subset \mathcal{A}$ is a field. If μ_1, μ_2 are σ-finite on $\mathcal{F}$, that is, there exist $A_1, A_2, \ldots \in F$ such that*

$$\bigcup_{n=1}^{\infty} A_n = \Omega \text{ and } \mu_i(A_n) < \infty \text{ for all } i = 1, 2,\ n = 1, 2, \ldots,$$

and

$$\mu_1(A) = \mu_2(A) \text{ for all } A \in \mathcal{F},$$

then

$$\mu_1(A) = \mu_2(A) \text{ for all } A \in \sigma(\mathcal{F})\,.$$

Proof of Theorem 2.4.2. Let $\mathcal{M}_0$ be the intersection of all monotone classes containing $\mathcal{F}$, that is, the smallest monotone class containing $\mathcal{F}$. Clearly, it suffices to show that

$$\sigma(\mathcal{F}) \subset \mathcal{M}_0\,.$$

To achieve that end, we shall show that $\mathcal{M}_0$ is a σ-field, which will be done via the following steps.

Step 1. If $A \in \mathcal{M}_0$, then $A^c \in \mathcal{M}_0$.

Proof of Step 1. Define

$$\mathcal{G} := \{A \in \mathcal{M}_0 : A^c \in \mathcal{M}_0\}\,.$$

Notice that $\mathcal{F} \subset \mathcal{G}$ because $\mathcal{F}$ is a field. Next observe that if $A_1, A_2, \ldots \in \mathcal{G}$ and $A_n \uparrow A$, then $A \in \mathcal{M}_0$ because $\mathcal{M}_0$ is closed under monotone unions. Further, $A_n^c \in \mathcal{M}_0$ and $A_n^c \downarrow A^c$. As $\mathcal{M}_0$ is closed under monotone intersection, $A^c \in \mathcal{M}_0$. Therefore, $A \in \mathcal{G}$, showing that $\mathcal{G}$ is closed under monotone union. Similarly, it can be shown that $\mathcal{G}$ is closed under monotone intersection. Thus, $\mathcal{G}$ is a monotone class containing $\mathcal{F}$. Hence, $\mathcal{G} \supset \mathcal{M}_0$, thereby proving Step 1. $\square$

Step 2. If $A \in \mathcal{F}$ and $B \in \mathcal{M}_0$, then $A \cup B \in \mathcal{M}_0$.

Proof. Define

$$\mathcal{H} := \{B \in \mathcal{M}_0 : A \cup B \in \mathcal{M}_0 \text{ for all } A \in \mathcal{F}\}.$$

By virtue of being a field, $\mathcal{F} \subset \mathcal{H}$. Routine verification will ensure that $\mathcal{H}$ is a monotone class, and hence contains $\mathcal{M}_0$. This proves Step 2. ☐

Step 3. If $A, B \in \mathcal{M}_0$, then $A \cup B \in \mathcal{M}_0$.

Proof. Define

$$\mathcal{I} := \{B \in \mathcal{M}_0 : A \cup B \in \mathcal{M}_0 \text{ for all } A \in \mathcal{M}_0\}.$$

By Step 2, it follows that $\mathcal{F} \subset \mathcal{I}$. Once again, similar ideas will ensure that $\mathcal{I}$ is a monotone class, and thus prove Step 3. ☐

Steps 1 and 3, along with the fact that $\mathcal{M}_0$ is a monotone class establishes that it is a σ-field, and thus completes the proof of the monotone class theorem. ☐

The statement of Theorem 2.4.3 should not be misinterpreted as the following: if μ, ν are σ-finite measures on $(\Omega, \sigma(\mathcal{F}))$, where $\mathcal{F}$ is a field, which agree on $\mathcal{F}$, then they agree on $\sigma(\mathcal{F})$. As shown by the following example, the claim just made is false.

Example 2.4.1. *Define measures* μ *and* ν *on* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *by*

$$\begin{aligned} \mu(A) &= \#(A \cap \mathbb{Q}), \\ \nu(A) &= 2\#(A \cap \mathbb{Q}), \end{aligned}$$

for all $A \in \mathcal{B}(\mathbb{R})$. *Let*

$$\mathcal{F} = \big\{\mathbb{R} \cap ((a_1, b_1] \cup \ldots \cup (a_n, b_n]) : -\infty \le a_1 < b_1 < \ldots < a_n < b_n \le \infty, \\ n = 0, 1, 2, \ldots\big\}.$$

Then $\mathcal{F}$ *is a field and* $\sigma(\mathcal{F}) = \mathcal{B}(\mathbb{R})$. *Further,* μ *and* ν *agree on* $\mathcal{F}$, *are* σ*-finite on* $\mathcal{B}(\mathbb{R})$, *and do not agree on* $\mathcal{B}(\mathbb{R})$.

2.5 Extension from semi-field to field

In practice, it is usually easier to define a set function on a “semi-field” and then extend it to the generated field.

Definition 2.5.1. *For $\Omega \neq \emptyset$, $\mathcal{S} \subset 2^\Omega$ is a* semi-field *on Ω if $\Omega \in \mathcal{S}$,*

$$A, B \in \mathcal{S} \text{ implies } A \cap B \in \mathcal{S},$$

and for all $A \in \mathcal{S}$,

$$A^c = A_1 \cup \ldots \cup A_n,$$

for some disjoint $A_1, \ldots, A_n \in \mathcal{S}$.

Example 2.5.1. *Let*

$$\mathcal{S} = \{(a, b] \cap \mathbb{R} : -\infty \leq a \leq b \leq \infty\} .$$

Then $\mathcal{S}$ is a semi-field on $\mathbb{R}$. It can be checked that $\sigma(\mathcal{S}) = \mathcal{B}(\mathbb{R})$.

Theorem 2.5.1. *If $\mathcal{S}$ is a semi-field on $\Omega \neq \emptyset$, then*

$$\mathcal{F} = \{A_1 \cup \ldots \cup A_n : A_1, \ldots, A_n \in \mathcal{S} \text{ are disjoint, } n = 0, 1, 2, \ldots\}$$

is the smallest field containing $\mathcal{S}$.

Proof. Trivially, any field containing $\mathcal{S}$ contains $\mathcal{F}$. Thus, all that has to be shown is $\mathcal{F}$ is a field. Since $\mathcal{S} \subset \mathcal{F}$, $\Omega \in \mathcal{F}$ is automatic.

Suppose $A, B \in \mathcal{F}$. Then

$$A = A_1 \cup \ldots \cup A_m \text{ and } B = B_1 \cup \ldots \cup B_n$$

where $A_1, \ldots, A_m \in \mathcal{S}$ are disjoint and so are $B_1, \ldots, B_n$. Thus

$$A \cap B = \bigcup_{i=1}^{m} \bigcup_{j=1}^{n} (A_i \cap B_j)$$

where $\{(A_i \cap B_j) : i = 1, \ldots, m,\ j = 1, \ldots, n\}$ is a collection of disjoint sets. Since $\mathcal{S}$ is a semi-field,

$$A_i \cap B_j \in \mathcal{S},\ i = 1, \ldots, m,\ j = 1, \ldots, n .$$

Hence $A \cap B \in \mathcal{F}$. That is, $\mathcal{F}$ is closed under finite intersections.

To show that $\mathcal{F}$ is closed under complements, let $A \in \mathcal{F}$. Then

$$A = A_1 \cup \ldots \cup A_n$$

for some disjoint $A_1, \ldots, A_n \in \mathcal{S}$. Since $A_i \in \mathcal{S}$ for $i = 1, \ldots, n$,

$$A_i^c = \bigcup_{j=1}^{k_i} B_{ij} \text{ for some disjoint } B_{i1}, \ldots, B_{ik_i} \in \mathcal{S} .$$

Thus

$$A^c = A_1^c \cap \ldots \cap A_n^c = \bigcup_{j_1=1}^{k_1} \ldots \bigcup_{j_n=1}^{k_n} \left(B_{1j_1} \cap \ldots \cap B_{nj_n}\right) . \tag{2.5.1}$$

Clearly,

$$B_{1j_1} \cap \ldots \cap B_{nj_n} \in \mathcal{S},\ (j_1, \ldots, j_n) \in \{1, \ldots, k_1\} \times \ldots \times \{1, \ldots, k_n\},$$

and

$$(B_{1i_1} \cap \ldots \cap B_{ni_n}) \cap (B_{1j_1} \cap \ldots \cap B_{nj_n}) = \emptyset \text{ if } (i_1, \ldots, i_n) \neq (j_1, \ldots, j_n) .$$

Hence (2.5.1) shows $A^c \in \mathcal{F}$. Since a collection of subsets, to which contains the whole space belongs and which is closed under complements and finite intersections, can be shown to be a field, $\mathcal{F}$ is a field, and hence the proof follows. ☐

Definition 2.5.2. *Suppose $\mathcal{S}$ is a semi-field on $\Omega \neq \emptyset$. A function $\mu : \mathcal{S} \to [0, \infty]$ is a* finitely additive set function *if $\mu(\emptyset) = 0$ and*

$$\mu(A_1 \cup \ldots \cup A_n) = \mu(A_1) + \ldots + \mu(A_n)$$

whenever $A_1, \ldots, A_n \in \mathcal{S}$ are disjoint such that $A_1 \cup \ldots \cup A_n \in \mathcal{S}$. If in addition,

$$\mu(A_1 \cup A_2 \cup \ldots) = \sum_{n=1}^{\infty} \mu(A_n)$$

whenever $A_1, A_2, \ldots \in \mathcal{S}$ are disjoint such that $A_1 \cup A_2 \cup \ldots \in \mathcal{S}$, then μ is a countably additive set function.

Theorem 2.5.2. *Suppose $\Omega \neq \emptyset$, $\mathcal{S}$ is a semi-field on Ω and μ is a finitely additive set function on $\mathcal{S}$. Let $\mathcal{F}$ be the field generated by $\mathcal{S}$. Define $\mu : \mathcal{F} \to [0, \infty]$ by*

$$\mu(A) = \sum_{i=1}^{k} \mu(A_i), \text{ if } A = A_1 \cup \ldots \cup A_k \text{ for disjoint } A_1, \ldots, A_k \in \mathcal{S} . \tag{2.5.2}$$

Then μ is well defined on $\mathcal{F}$, that is, different representations of A lead to the same definition of $\mu(A)$, definition of μ remains unchanged on $\mathcal{S}$ and μ is finitely additive on $\mathcal{F}$. If in addition, μ is countably additive on $\mathcal{S}$, then so it is on $\mathcal{F}$.

Proof. Assume μ is finitely additive on $\mathcal{S}$. The first step is to show that the right hand side of (2.5.2) remains invariant under the choice of $A_1, \ldots, A_k$. In other words, if for some $A \in \mathcal{F}$,

$$A_1 \cup \ldots \cup A_m = A = B_1 \cup \ldots \cup B_n,$$

where $A_1, \ldots, A_m$ are disjoint sets in $\mathcal{S}$ and so are $B_1, \ldots, B_n$, then

$$\sum_{i=1}^{m} \mu(A_i) = \sum_{j=1}^{n} \mu(B_j) \,. \tag{2.5.3}$$

For a fixed $i = 1, \ldots, m$,

$$A_i = A_i \cap A = A_i \cap (B_1 \cup \ldots \cup B_n) = \bigcup_{j=1}^{n} (A_i \cap B_j) \,.$$

Since $B_1, \ldots, B_n$ are disjoint, so are $A_i \cap B_1, \ldots, A_i \cap B_n$. Finite additivity of μ shows

$$\mu(A_i) = \sum_{j=1}^{n} \mu(A_i \cap B_j),$$

and hence

$$\sum_{i=1}^{m} \mu(A_i) = \sum_{i=1}^{m} \sum_{j=1}^{n} \mu(A_i \cap B_j) \,.$$

A similar argument shows that

$$\sum_{j=1}^{n} \mu(B_j) = \sum_{j=1}^{n} \sum_{i=1}^{m} \mu(A_i \cap B_j) \,;$$

(2.5.3) follows by comparing the above two equalities. In other words, μ is well defined by (2.5.2). A trivial consequence is that (2.5.2) keeps μ unchanged on $\mathcal{S}$.

Finite additivity of μ on $\mathcal{F}$ would follow if it shown that

$$\mu(A \cup B) = \mu(A) + \mu(B) \text{ for disjoint } A, B \in \mathcal{F} \,.$$

Fix $A, B \in \mathcal{F}$ disjoint. Then

$$A = A_1 \cup \ldots \cup A_m \text{ and } B = B_1 \cup \ldots \cup B_n$$

for disjoint $A_1, \ldots, A_m, B_1, \ldots, B_n \in \mathcal{S}$. It follows from (2.5.2) that

$$\mu(A \cup B) = \sum_{i=1}^{m} \mu(A_i) + \sum_{j=1}^{n} \mu(B_j) = \mu(A) + \mu(B),$$

establishing μ is finitely additive on $\mathcal{F}$.

For the final claim, suppose μ is countably additive on $\mathcal{S}$. Suppose that $A_1, A_2, \ldots \in \mathcal{F}$ and

$$\bigcup_{n=1}^{\infty} A_n = A_\infty \in \mathcal{F} \,.$$

Then for $n = 1, 2, \ldots, \infty$, there exist disjoint $B_{n1}, \ldots, B_{nk_n} \in \mathcal{S}$ such that

$$A_n = B_{n1} \cup \ldots \cup B_{nk_n},$$

where $k_1, k_2, \ldots, k_\infty \in \mathbb{N}$. Thus

$$\mu(A_\infty) = \sum_{i=1}^{k_\infty} \mu\left(B_{\infty i}\right) . \tag{2.5.4}$$

For $i = 1, \ldots, k_\infty$,

$$\begin{aligned} B_{\infty i} &= B_{\infty i} \cap A_\infty \\ &= B_{\infty i} \cap \left(\bigcup_{n=1}^{\infty} \bigcup_{j=1}^{k_n} B_{nj} \right) \\ &= \bigcup_{n=1}^{\infty} \bigcup_{j=1}^{k_n} \left(B_{\infty i} \cap B_{nj}\right) . \end{aligned}$$

Since $B_{\infty i} \cap B_{nj} \in S$ for all $n = 1, 2, \ldots$ and $j = 1, \ldots, k_n$, and the right hand side above is a disjoint union, countable additivity of μ on $\mathcal{S}$ shows that

$$\mu(B_{\infty i}) = \sum_{n=1}^{\infty} \sum_{j=1}^{k_n} \mu(B_{\infty i} \cap B_{nj}) .$$

Invoking (2.5.4) in conjunction with the above, it follows that

$$\begin{aligned} \mu(A_\infty) &= \sum_{i=1}^{k_\infty} \sum_{n=1}^{\infty} \sum_{j=1}^{k_n} \mu(B_{\infty i} \cap B_{nj}) \\ &= \sum_{n=1}^{\infty} \sum_{i=1}^{k_\infty} \sum_{j=1}^{k_n} \mu(B_{\infty i} \cap B_{nj}) \\ &= \sum_{n=1}^{\infty} \mu(A_n), \end{aligned}$$

where the last line again uses (2.5.2) along with the observation that

$$A_n = A_\infty \cap A_n = \bigcup_{i=1}^{k_\infty} \bigcup_{j=1}^{k_n} (B_{\infty i} \cap B_{nj}),$$

and that the extreme right hand side is a disjoint union. This shows μ is countably additive on $\mathcal{F}$ and hence completes the proof. $\square$

The following is a trivial consequence of Theorems 2.3.1 and 2.5.2.

Corollary 2.5.1. *If $\Omega \neq \emptyset$, $\mathcal{S}$ is a semi-field on Ω and $\mu : \mathcal{S} \to [0,\infty]$ is countably additive, then μ can be extended to a measure on $\big(\Omega, \sigma(\mathcal{S})\big)$.*

Theorem 2.5.3. *If $\Omega \neq \emptyset$, $\mathcal{F}$ is a field on Ω, and μ is a finitely additive set function on $\mathcal{F}$, then μ is*

(a) monotone, that is, $\mu(A) \le \mu(B)$ for $A, B \in \mathcal{F}$ with $A \subset B$,

(b) finitely subadditive, that is,

$$\mu(A_1 \cup \ldots \cup A_n) \le \mu(A_1) + \ldots + \mu(A_n),\ A_1, \ldots, A_n \in \mathcal{F},$$

(c) and countably superadditive, that is,

$$\mu(A) \ge \sum_{n=1}^{\infty} \mu(A_n) \text{ if } A_1, A_2, \ldots \in \mathcal{F} \text{ disjoint, } A = \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}\,.$$

Proof. The first two claims are trivial. For (c), use (a) to argue for a finite N,

$$\mu(A) \ge \mu\left(A_1 \cup \ldots \cup A_N\right) = \sum_{n=1}^{N} \mu(A_n)\,.$$

Letting $N \to \infty$, (c) follows, which completes the proof. □

Regarding the question of uniqueness, the following is a trivial consequence of Theorems 2.4.3 and 2.5.1.

Theorem 2.5.4. *Suppose $(\Omega, \mathcal{A})$ is a measurable space on which, μ_1, μ_2 are measures. Assume $\mathcal{S} \subset \mathcal{A}$ is a semi-field. If μ_1, μ_2 are σ-finite on $\mathcal{S}$, that is, there exist $A_1, A_2, \ldots \in S$ such that*

$$\bigcup_{n=1}^{\infty} A_n = \Omega \text{ and } \mu_i(A_n) < \infty \text{ for all } i = 1, 2,\ n = 1, 2, \ldots,$$

and

$$\mu_1(A) = \mu_2(A) \text{ for all } A \in \mathcal{S},$$

then

$$\mu_1(A) = \mu_2(A) \text{ for all } A \in \sigma(\mathcal{S})\,.$$

Proof. Exercise. □

2.6 Riemann-Stieltjes measure on $\mathbb{R}$

Definition 2.6.1. *Given a function* $F:\mathbb{R}\to\mathbb{R}$*, a measure* μ *on* $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ *is the* Riemann-Stieltjes measure corresponding to F *if*

$$\mu\big((a,b]\big)=F(b)-F(a)\text{ for all }-\infty<a\le b<\infty\,. \tag{2.6.1}$$

Due to its importance, the following is a fundamental result in measure theory.

Theorem 2.6.1. *For any non-decreasing right continuous function* $F:\mathbb{R}\to\mathbb{R}$*, a unique Riemann-Stieltjes measure* μ *corresponding to* F *exists.*

Proof. Define

$$F(\infty)=\lim_{x\to\infty}F(x),$$

and

$$F(-\infty)=\lim_{x\to-\infty}F(x),$$

both of which exist because F is monotone. Let $\mathcal{S}$ be as in Example 2.5.1. Define $\mu:\mathcal{S}\to[0,\infty]$ by

$$\mu(A)=\begin{cases}F(b)-F(a), & A=(a,b]\cap\mathbb{R}\text{ for some }-\infty\le a<b\le\infty,\\ 0, & \text{otherwise, that is, if }A=\emptyset\,;\end{cases}$$

$F(b)-F(a)$ is defined whenever $a<b$ because then neither $F(b)$ nor $-F(a)$ equals $-\infty$ (the sum of two quantities in $[-\infty,\infty]$ is undefined if and only if one of them is ∞ and the other one $-\infty$).

The first step is to show that μ is finitely additive on $\mathcal{S}$. Let $A_1,\ldots,A_n\in\mathcal{S}$ be disjoint such that

$$A=A_1\cup\ldots\cup A_n\in\mathcal{S}\,.$$

Without loss of generality, assume these are non-empty, that is,

$$A_i=(a_i,b_i]\cap\mathbb{R}\text{ for some }-\infty\le a_i<b_i\le\infty\,.$$

Since $A_i\cap A_j=\emptyset$ for all $1\le i<j\le n$, either $b_i\le a_j$ or $b_j\le a_i$. By a relabelling, it can be assumed without loss of generality that

$$a_1<b_1\le a_2<\ldots<b_{n-1}\le a_n<b_n\,.$$

Since $A\in\mathcal{S}$, it is necessary that $A=(a_1,b_n]\cap\mathbb{R}$ and hence $b_1=a_2,\ldots,$ $b_{n-1}=a_n$. Thus

$$\begin{aligned}\sum_{i=1}^n\mu(A_i)&=\sum_{i=1}^n\big(F(b_i)-F(a_i)\big)\\ (\text{because }b_1=a_2,\ldots,b_{n-1}=a_n)&=F(b_n)-F(a_1)\\ &=\mu(A),\end{aligned}$$

showing μ is finitely additive on $\mathcal{S}$.

Let $\mathcal{F}$ be the field generated by $\mathcal{S}$. Extend μ to $\mathcal{F}$ by (2.5.2). Theorem 2.5.2 shows μ is finitely additive on $\mathcal{F}$. Theorem 2.5.3 tells us μ is monotone, finitely subadditive and countably superadditive on $\mathcal{F}$.

Our next task is to show μ is countably additive on $\mathcal{S}$. Let $A_1, A_2, \ldots \in \mathcal{S}$ be disjoint such that

$$\bigcup_{n=1}^{\infty} A_n = A \in \mathcal{S}\,.$$

Countable superadditivity of μ on $\mathcal{F}$ implies

$$\mu(A) \geq \sum_{n=1}^{\infty} \mu(A_n)\,.$$

Once it is shown that

$$\mu(A) \leq \varepsilon + \sum_{n=1}^{\infty} \mu(A_n) \text{ for all } \varepsilon > 0, \tag{2.6.2}$$

countable additivity of μ on $\mathcal{S}$ would follow.

Fix $\varepsilon > 0$. Suppose $A_n = (a_n, b_n] \cap \mathbb{R}$ for some $a_n < b_n$. Let $\delta_n > 0$ be such that

$$F(b_n + \delta_n) \leq \varepsilon 2^{-n} + F(b_n),\ n \geq 1,$$

which exists if $b_n < \infty$ by right continuity of F, and trivially holds for any $\delta_n > 0$ if $b_n = \infty$.

Suppose $A = (a, b] \subset \mathbb{R}$ for some $a < b$. The proof of (2.6.2) will be given separately for the cases $b < \infty$ and $b = \infty$. First assume $b < \infty$. Let $a' \in (a, b]$. In this case, $-\infty \leq a < a' \leq b < \infty$. Since

$$[a', b] \subset (a, b] = \bigcup_{n=1}^{\infty} \big((a_n, b_n] \cap \mathbb{R}\big) \subset \bigcup_{n=1}^{\infty} (a_n, b_n + \delta_n), \tag{2.6.3}$$

the Heine-Borel theorem implies

$$[a', b] \subset \bigcup_{n=1}^{N} (a_n, b_n + \delta_n)$$

for some finite N. Thus

$$(a', b] \subset [a', b] \subset \bigcup_{n=1}^{N} (a_n, b_n + \delta_n) \subset \bigcup_{n=1}^{N} \big((a_n, b_n + \delta_n] \cap \mathbb{R}\big)\,.$$

Monotonicity of μ on $\mathcal{F}$ implies

$$\begin{aligned}
\mu\big((a',b]\big) &\le \mu\left(\bigcup_{n=1}^{N}\big((a_n,b_n+\delta_n]\cap\mathbb{R}\big)\right)\\
\text{(finite subadditiviety on } \mathcal{F}) &\le \sum_{n=1}^{N}\mu\big((a_n,b_n+\delta_n]\cap\mathbb{R}\big)\\
&= \sum_{n=1}^{N}\big(F(b_n+\delta_n)-F(a_n)\big)\\
&\le \sum_{n=1}^{\infty}\big(\varepsilon 2^{-n}+F(b_n)-F(a_n)\big)\\
&= \varepsilon+\sum_{n=1}^{\infty}\mu(A_n)\,.
\end{aligned}$$

That is,

$$F(b)-F(a')\le\varepsilon+\sum_{n=1}^{\infty}\mu(A_n)\,.$$

Since

$$\lim_{a'\downarrow a}F(a')=F(a), \tag{2.6.4}$$

which follows from right continuity of F is $a>-\infty$ and the definition if $a=-\infty$, (2.6.2) follows for the case $b<\infty$.

To prove (2.6.2) in the case $b=\infty$, fix $a<a'\le b'<b=\infty$. The arguments from (2.6.3) to (2.6.4) with b replaced by b', using the Heine-Borel theorem, imply

$$F(b')-F(a)\le\varepsilon+\sum_{n=1}^{\infty}\mu(A_n)\,.$$

Letting $b'\uparrow\infty$, (2.6.2) follows for the case $b=\infty$.

As argued before, (2.6.2) shows μ is countably additive on $\mathcal{S}$. Invoking Corollary 2.5.1, μ can be extended to a measure on $(\mathbb{R},\sigma(\mathcal{S}))$. Clearly, $\sigma(\mathcal{S})=\mathcal{B}(\mathbb{R})$, and hence μ is a Riemann-Stieltjes measure corresponding to F.

For uniqueness, suppose μ_1 and μ_2 are Riemann-Stieltjes measure corresponding to F. Then

$$\mu_1\big((a,b]\big)=\mu_2\big((a,b]\big)\text{ for all }-\infty<a\le b<\infty\,.$$

Continuity from below implies

$$\mu_1(A)=\mu_2(A)\text{ for all }A\in\mathcal{S}\,.$$

Further,

$$\mathbb{R} = \bigcup_{n\in\mathbb{Z}} (n, n+1] \text{ and } \mu_i\big((n, n+1]\big) < \infty \text{ for all } n \in \mathbb{Z},\ i = 1, 2\,.$$

Theorem 2.5.4 implies μ_1 and μ_2 agree on $\sigma(\mathcal{S})$ which is $\mathcal{B}(\mathbb{R})$. Thus uniqueness follows, which completes the proof. ☐

Theorem 2.6.2. *Given a function $F : \mathbb{R} \to \mathbb{R}$, a Riemann-Stieltjes measure corresponding to F exists if and only if F is non-decreasing and right continuous.*

Proof. The "if" part is precisely the claim of Theorem 2.6.1. For the converse part, assume a Riemann-Stieltjes measure μ corresponding to F exists. Then for $-\infty < a < b < \infty$,

$$F(b) - F(a) = \mu\big((a, b]\big) \geq 0,$$

showing that F is non-decreasing. For right continuity, assume that $x_n \downarrow x_\infty$ for real numbers $x_1, \ldots, x_\infty$. Then

$$(x_\infty, x_n] \downarrow \emptyset\,.$$

Since $\mu\big((x_\infty, x_1]\big) = F(x_1) - F(x_\infty) < \infty$, continuity from above implies

$$\mu\big((x_\infty, x_n]\big) \downarrow 0\,.$$

In other words,

$$\lim_{n\to\infty} F(x_n) = F(x_\infty)\,.$$

As this holds for any $x_n \downarrow x_\infty$, F is right continuous at x_∞. Hence the "only if" part follows, which completes the proof. ☐

Definition 2.6.2. *The Riemann-Stieltjes measure corresponding to F, when F is the identity function, that is,*

$$F(x) = x \text{ for all } x \in \mathbb{R},$$

is the Lebesgue measure *on $\mathbb{R}$.*

Exercise 2.6.1.

1. *Show that the Lebesgue measure λ on $\mathbb{R}$ is translation invariant, that is,*

   $$\lambda(A + x) = \lambda(A) \text{ for all } x \in \mathbb{R},\ A \in \mathcal{B}(\mathbb{R}),$$

   where $A + x = \{a + x : a \in A\}$.

2. *Hence or otherwise, prove that the Vitali set V, as in* (2.1.3), *is not a Borel set.*

Definition 2.6.3. *A measure μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is a* Radon measure *if $\mu(K) < \infty$ for every compact set $K \subset \mathbb{R}$.*

Exercise 2.6.2. *Suppose μ is a measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Show that there exists a non-decreasing right continuous $F : \mathbb{R} \to \mathbb{R}$ such that μ is the Riemann-Stieltjes measure corresponding to F if and only if μ is a Radon measure.*

3 Integration

Once measure has been understood, the next step is to develop the theory of integration. The first step towards that is to study those functions for which integration can possibly be defined, which are so-called "measurable functions".

3.1 Measurable functions

Definition 3.1.1. *Suppose $(\Omega_1, \mathcal{A}_1)$ and $(\Omega_2, \mathcal{A}_2)$ are measurable spaces. A function $f : \Omega_1 \to \Omega_2$ is* measurable *with respect to $\mathcal{A}_1/\mathcal{A}_2$ if*

$$f^{-1}(A) \in \mathcal{A}_1 \text{ for all } A \in \mathcal{A}_2,$$

where $f^{-1}(A) = \{\omega \in \Omega_1 : f(\omega) \in A\}$. Whenever either $\mathcal{A}_1$ or $\mathcal{A}_2$ is obvious from the context, it will be suppressed in the mention of measurability.

Theorem 3.1.1. *Suppose $\Omega_1 \neq \emptyset$ and $(\Omega_2, \mathcal{A}_2)$ is a measurable space. Then*

$$\left\{f^{-1}A : A \in \mathcal{A}_2\right\}$$

is a σ-field.

Proof. Let

$$\mathcal{A}_1 = \left\{f^{-1}A : A \in \mathcal{A}_2\right\}.$$

Clearly, $\Omega_1 = f^{-1}\Omega_2 \in \mathcal{A}_1$. It is easy to check that for any collection $(A_\alpha : \alpha \in I)$ of subsets of Ω_2,

$$f^{-1}\left(\bigcup_{\alpha \in I} A_\alpha\right) = \bigcup_{\alpha \in I} f^{-1}A_\alpha \,. \tag{3.1.1}$$

Hence if $B_1, B_2, \ldots \in \mathcal{A}_1$, that is, $B_n = f^{-1}A_n$ for some $A_n \in \mathcal{A}_2$, then

$$B_1 \cup B_2 \cup \ldots = f^{-1}(A_1 \cup A_2 \cup \ldots) \in \mathcal{A}_1 \text{ because } A_1 \cup A_2 \cup \ldots \in \mathcal{A}_2 \,.$$

Finally,

$$\left(f^{-1}A\right)^c = f^{-1}(A^c) \text{ for all } A \subset \Omega_2, \tag{3.1.2}$$

which shows $\mathcal{A}_1$ is closed under complements. Thus $\mathcal{A}_1$ is a σ-field, which completes the proof. ☐

Theorem 3.1.2. *Suppose* $(\Omega_1, \mathcal{A}_1)$ *and* $(\Omega_2, \mathcal{A}_2)$ *are measurable spaces and* $f : \Omega_1 \to \Omega_2$ *is a function. If* $\mathcal{G} \subset \mathcal{A}_2$ *is such that* $\sigma(\mathcal{G}) = \mathcal{A}_2$ *and*

$$f^{-1}A \in \mathcal{A}_1 \text{ for all } A \in \mathcal{G},$$

then f *is measurable.*

Proof. Let

$$\mathcal{F} = \left\{A \subset \Omega_2 : f^{-1}A \in \mathcal{A}_1\right\}.$$

The hypothesis implies $\mathcal{G} \subset \mathcal{F}$. Clearly $\Omega_2 \in \mathcal{F}$ because $\mathcal{A}_1$ is a σ-field and hence $f^{-1}\Omega_2 = \Omega_1 \in \mathcal{A}_1$. Next (3.1.1) and (3.1.2) show $\mathcal{F}$ is a σ-field. Thus $\mathcal{F} \supset \sigma(\mathcal{G}) = \mathcal{A}_2$. Hence the proof follows. ☐

Exercise 3.1.1. *Show that*

$$\mathcal{B}(\mathbb{R}) = \sigma\left(\{(-\infty, x) : x \in \mathbb{R}\}\right) = \sigma\left(\{(-\infty, x] : x \in \mathbb{R}\}\right).$$

Theorem 3.1.3. *Suppose* $(\Omega, \mathcal{A})$ *is a measurable space and* f, g *are* $\mathcal{A}/\mathcal{B}(\mathbb{R})$*-measurable functions from* Ω *to* $\mathbb{R}$*. Then* $f + g$ *is* $\mathcal{A}/\mathcal{B}(\mathbb{R})$*-measurable.*

Proof. In view of Exercise 3.1.1, it suffices to show that

$$(f+g)^{-1}(-\infty, x) \in \mathcal{A} \text{ for all } x \in \mathbb{R}.$$

Fix $x \in \mathbb{R}$. We claim that

$$(f+g)^{-1}(-\infty, x) = \bigcup_{r \in \mathbb{Q}} \left(f^{-1}(-\infty, r)\right) \cap \left(g^{-1}(-\infty, x - r)\right).$$

That the right hand side is a subset of the left hand side is obvious. For the reverse inclusion, fix ω in the left hand side, that is,

$$f(\omega) + g(\omega) < x.$$

Then $f(\omega) < x - g(\omega)$. Thus there exists $r \in \mathbb{Q}$ such that

$$f(\omega) < r < x - g(\omega).$$

Hence $\omega \in \left(f^{-1}(-\infty, r)\right) \cap \left(g^{-1}(-\infty, x - r)\right)$. This shows the claimed set theoretic equality and hence the proof follows. ☐

Exercise 3.1.2. *If* $(\Omega, \mathcal{A})$ *is a measurable space and* $f : \Omega \to \mathbb{R}$ *is measurable, show that for any* $\alpha \in \mathbb{R}$, αf *is measurable.*

Definition 3.1.2. *A function* $f : \mathbb{R} \to \mathbb{R}$ *is a* Borel function *if it is* $\mathcal{B}(\mathbb{R})/\mathcal{B}(\mathbb{R})$*-measurable.*

Theorem 3.1.4. *1. A continuous function is a Borel function.*

2. A monotone function is a Borel function.

Proof. 1. If $f : \mathbb{R} \to \mathbb{R}$ is continuous, then $f^{-1}U$ is an open set for every open set $U \subset \mathbb{R}$. Since

$$\mathcal{B}(\mathbb{R}) = \sigma\left(\{U \subset \mathbb{R} : U \text{ is open}\}\right),$$

f is Borel.

2. Suppose $f : \mathbb{R} \to \mathbb{R}$ is non-decreasing. Fix $x \in \mathbb{R}$ and let

$$\alpha = \sup\left(f^{-1}(-\infty, x]\right) .$$

It is immediate that $f^{-1}(-\infty, x]$ is either $(-\infty.\alpha)$ or $(-\infty, \alpha]$. In both cases,

$$f^{-1}(-\infty, x] \in \mathcal{B}(\mathbb{R}) .$$

Since this holds for all x, f is Borel. A similar argument holds for a non-increasing function.

□

Theorem 3.1.5. *Suppose $(\Omega_i, \mathcal{A}_i)$ is a measurable space for $i = 1, 2, 3$. If $f : \Omega_1 \to \Omega_2$ and $g : \Omega_2 \to \Omega_3$ are measurable, then $g \circ f : \Omega_1 \to \Omega_3$ is measurable.*

Proof. Follows from the definition. □

Theorem 3.1.6. *Suppose $(\Omega, \mathcal{A})$ is a measurable space and f, g are measurable functions from Ω to $\mathbb{R}$. Then fg is measurable.*

Proof. Write

$$fg = \left(\frac{f+g}{2}\right)^2 - \left(\frac{f-g}{2}\right)^2 .$$

Theorem 3.1.3 shows $(f+g)/2$ and $(f-g)/2$ are measurable. The function $x \mapsto x^2$ is Borel by Theorem 3.1.4. Theorem 3.1.5 shows $\left((f \pm g)/2\right)^2$ is measurable. The proof follows by applying Theorem 3.1.3 once more. □

Theorem 3.1.7. *If f, g are measurable functions from some measurable space $(\Omega, \mathcal{A})$ to $\mathbb{R}$, then $f \vee g$ and $f \wedge g$ are measurable.*

Proof. Exercise. □

3.2 Integration of non-negative functions

Henceforth, $(\Omega, \mathcal{A}, \mu)$, where $\mu(\Omega) > 0$, will be the underlying measure space. Unless mentioned otherwise, functions talked about are $\mathcal{A}$-measurable. The theory of integration is first developed for all non-negative measurable functions on Ω which possibly take the value infinity. To that end, define

$$\overline{\mathbb{R}} = [-\infty, \infty]$$

and

$$\mathcal{B}(\overline{\mathbb{R}}) = \sigma\left(\mathcal{B}(\mathbb{R}) \cup \left\{\{-\infty\}, \{\infty\}\right\}\right) .$$

That is, $\mathcal{B}(\overline{\mathbb{R}})$ is the smallest σ-field on $\overline{\mathbb{R}}$ which contains the singleton sets $\{-\infty\}$ and $\{\infty\}$, in addition to every Borel subset of $\mathbb{R}$.

Exercise 3.2.1. *Show that*

$$\mathcal{B}(\overline{\mathbb{R}}) = \sigma\left(\{[-\infty, x) : x \in \mathbb{R}\}\right) .$$

Definition 3.2.1. *A measurable function $s : \Omega \to \overline{\mathbb{R}}$ is a* simple function *if range of s is a finite set, that is,*

$$\#s(\Omega) < \infty .$$

First we shall define integral of simple functions, for which, the conventions of multiplication of two numbers which are possibly infinite have to be set. Define

$$x \cdot \infty = \infty \cdot x = \begin{cases} -\infty, & \text{if } -\infty \le x < 0, \\ 0, & \text{if } x = 0, \\ \infty, & \text{if } 0 < x \le \infty, \end{cases}$$

and

$$x \cdot (-\infty) = (-\infty) \cdot x = \begin{cases} \infty, & \text{if } -\infty \le x < 0, \\ 0, & \text{if } x = 0, \\ -\infty, & \text{if } 0 < x \le \infty . \end{cases}$$

It is worth emphasizing that $\pm\infty \cdot 0$ has been defined to be 0.

Definition 3.2.2. *If $s : \Omega \to [0, \infty]$ is a simple function whose range is $\{\alpha_1, \ldots \ldots, \alpha_n\}$, where $\alpha_1, \ldots, \alpha_n$ are distinct by convention, the integral of s with respect to μ, denoted by $\int s\, d\mu$, is defined by*

$$\int s\, d\mu = \sum_{i=1}^{n} \alpha_i \mu\left(s^{-1}\{\alpha_i\}\right) . \tag{3.2.1}$$

Theorem 3.2.1. *If $s = \sum_{i=1}^{n} \beta_i \mathbf{1}_{A_i}$, where $A_1, \ldots, A_n \in \mathcal{A}$ are disjoint and $\beta_1, \ldots, \beta_n \in [0, \infty]$, then*

$$\int s\, d\mu = \sum_{i=1}^{n} \beta_i \mu(A_i) .$$

Proof. Zero times anything, including ∞, is zero. Hence it can be assumed without loss of generality that $A_i \neq \emptyset$ for $i = 1, \ldots, n$. For $\omega \in A_i$, which exists as $A_i \neq \emptyset$,

$$s(\omega) = \beta_i,$$

because $A_1, \ldots, A_n$ are disjoint. Thus $\beta_1, \ldots, \beta_n \in s(\Omega)$. Further, for all $\alpha \in s(\Omega)$, either $\alpha = 0$ or $\alpha = \beta_i$ for some i. In the latter case,

$$s^{-1}\{\alpha\} = \bigcup_{i:\beta_i=\alpha} A_i,$$

and hence for $\alpha \in s(\Omega) \setminus \{0\}$,

$$\mu\left(s^{-1}\{\alpha\}\right) = \sum_{i:\beta_i=\alpha} \mu(A_i)\,. \tag{3.2.2}$$

The definition implies

$$\begin{aligned}
\int s\,d\mu &= \sum_{\alpha\in s(\Omega)} \alpha\mu\left(s^{-1}\{\alpha\}\right) \\
&= \sum_{\alpha\in s(\Omega)\setminus\{0\}} \alpha\mu\left(s^{-1}\{\alpha\}\right) \\
\big(\text{by (3.2.2)}\big) &= \sum_{\alpha\in s(\Omega)\setminus\{0\}} \alpha \sum_{i:\beta_i=\alpha} \mu(A_i) \\
&= \sum_{\alpha\in s(\Omega)\setminus\{0\}} \sum_{i:\beta_i=\alpha} \beta_i\mu(A_i) \\
&= \sum_{i:\beta_i\neq 0} \beta_i\mu(A_i) \\
&= \sum_{i=1}^{n} \beta_i\mu(A_i)\,.
\end{aligned}$$

This completes the proof. ☐

Theorem 3.2.2. *If s, t are simple functions with $0 \leq s \leq t$, then*

$$\int s\,d\mu \leq \int t\,d\mu\,.$$

Proof. Let

$$\left\{s^{-1}\{\alpha\} \cap t^{-1}\{\beta\} : \alpha \in s(\Omega),\, \beta \in t(\Omega)\right\} = \{A_1, \ldots, A_n\}\,.$$

For $i = 1, \ldots, n$, if $A_i \neq \emptyset$, fix $\omega \in A_i$ and denote

$$\alpha_i = s(\omega),\ \beta_i = t(\omega),$$

and let $\alpha_i = \beta_i = 0$ if $A_i = \emptyset$. Thus $\alpha_i \le \beta_i$ for all i, $A_1, \ldots, A_n$ are disjoint sets in $\mathcal{A}$ and

$$s = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i},\ t = \sum_{i=1}^n \beta_i \mathbf{1}_{A_i}\,.$$

Theorem 3.2.1 implies

$$\begin{aligned} \int s\, d\mu &= \sum_{i=1}^n \alpha_i \mu(A_i) \\ &\le \sum_{i=1}^n \beta_i \mu(A_i) \\ &= \int t\, d\mu, \end{aligned}$$

which completes the proof. □

Definition 3.2.3. *For a measurable function* $f : \Omega \to [0, \infty]$, *define*

$$\int f\, d\mu = \sup\left\{\int s\, d\mu : \ 0 \le s \le f,\ s \textit{ is a simple function}\right\}\,.$$

Theorem 3.2.2 ensures that for a simple function $t \ge 0$, the above definition is consistent with (3.2.1), that is,

$$\sum_{\alpha \in t(\Omega)} \alpha\mu\left(t^{-1}\{\alpha\}\right) = \sup\left\{\int s\, d\mu : \ 0 \le s \le t,\ s \text{ is a simple function}\right\}\,.$$

The following theorem is a trivial consequence of the definition given above.

Theorem 3.2.3. *For measurable* f, g *with* $0 \le f \le g$,

$$\int f\, d\mu \le \int g\, d\mu\,.$$

Proof. Exercise. □

Theorem 3.2.4. *Suppose* $f_1, f_2, \ldots$ *are measurable functions from* $\Omega \to \overline{\mathbb{R}}$. *Then*

$$g = \inf_n f_n \ \textit{and}\ h = \sup_n f_n$$

are measurable.

Proof. Follows trivially from the observation that

$$[g < \alpha] = \bigcup_{n=1}^\infty [f_n < \alpha]$$

and Exc 3.2.1 and likewise for h. □

Exercise 3.2.2. *For measurable functions $f_1, f_2, \ldots$ from Ω to $\overline{\mathbb{R}}$, show the following are measurable:*

1. *$f_1 + f_2$, if it is defined,*
2. *$f_1 f_2$,*
3. *$\liminf_{n\to\infty} f_n$,*
4. *$\limsup_{n\to\infty} f_n$,*
5. *$\lim_{n\to\infty} f_n$, if it exists.*

The next theorem is of utmost importance in measure theory. As mentioned earlier, 'MCT' will refer to the following result and not the monotone class theorem (Theorem 2.4.2).

Theorem 3.2.5 (Monotone convergence theorem (MCT))**.** *If $f_n \geq 0$ and $f_n \uparrow f$, then*

$$\int f_n \, d\mu \uparrow \int f \, d\mu \,,$$

f being measurable by Theorem 3.2.4.

The following exercise will be used for proving the above theorem.

Exercise 3.2.3. *If $\alpha_1, \ldots, \alpha_\infty, \beta_1, \ldots, \beta_\infty$ are such that $0 \leq \alpha_n \uparrow \alpha_\infty$ and $0 \leq \beta_n \uparrow \beta_\infty$, then*

$$\alpha_n \beta_n \uparrow \alpha_\infty \beta_\infty \,.$$

Proof of Theorem 3.2.5 (MCT). Theorem 3.2.3 implies that

$$\int f_1 \, d\mu \leq \int f_2 \, d\mu \leq \ldots \leq \int f \, d\mu \,.$$

Hence

$$\int f \, d\mu \geq \lim_{n\to\infty} \int f_n \, d\mu, \text{ which exists.}$$

To complete the proof, it suffices to show that for all

$$\alpha < \int f \, d\mu,$$

there exists n for which

$$\int f_n \, d\mu \geq \alpha \,. \tag{3.2.3}$$

Fix α as above. The definition of integral implies there exists a simple function s such that $0 \leq s \leq f$ and $\int s \, d\mu > \alpha$. Write

$$s = \sum_{i=1}^{k} \beta_i \mathbf{1}_{A_i}$$

where $A_1, \ldots, A_k \in \mathcal{A}$ are disjoint and $\beta_1, \ldots, \beta_k > 0$. Since

$$\begin{aligned} \alpha &< \int s \, d\mu \\ \text{(Theorem 3.2.1)} &= \sum_{i=1}^{k} \beta_i \mu(A_i) \\ &= \lim_{\gamma_1 \uparrow \beta_1, \ldots, \gamma_k \uparrow \beta_k} \sum_{i=1}^{k} \gamma_i \mu(A_i), \end{aligned}$$

Exc 3.2.3 implying the last line, there exist $0 \le \gamma_1 < \beta_1, \ldots, 0 \le \gamma_k < \beta_k$ such that

$$\sum_{i=1}^{k} \gamma_i \mu(A_i) > \alpha \,. \tag{3.2.4}$$

Let

$$A_{ni} = \{\omega \in A_i : f_n(\omega) > \gamma_i\}, \; n = 1, 2, \ldots, \; \text{ and } i = 1, \ldots, k \,.$$

If $\omega \in A_i$, then

$$\lim_{n \to \infty} f_n(\omega) = f(\omega) \ge s(\omega) = \beta_i > \gamma_i \,.$$

Hence

$$A_{ni} \uparrow A_i, \; i = 1. \ldots, k \,.$$

Continuity from below implies

$$\mu\left(A_{ni}\right) \uparrow \mu(A_i), \; n \to \infty \,.$$

Exc 3.2.3 shows

$$\lim_{n \to \infty} \sum_{i=1}^{k} \gamma_i \mu\left(A_{ni}\right) = \sum_{i=1}^{k} \gamma_i \mu(A_i) > \alpha,$$

(3.2.4) implying the inequality.

Thus there exists n for which

$$\begin{aligned} \alpha &< \sum_{i=1}^{k} \gamma_i \mu(A_{ni}) \\ \text{(Theorem 3.2.1)} &= \int \left(\sum_{i=1}^{k} \gamma_i \mathbf{1}_{A_{ni}} \right) d\mu \,. \end{aligned}$$

The definition of A_{ni} implies

$$\sum_{i=1}^{k} \gamma_i \mathbf{1}_{A_{ni}} \le f_n,$$

which in conjunction with the above shows

$$\alpha \le \int f_n \, d\mu \, .$$

This show (3.2.3) from which the proof follows. □

The following is an immediate consequence of the MCT.

Theorem 3.2.6. *For $f, g \ge 0$ measurable and $\alpha \in [0, \infty]$,*

$$\int (f+g) \, d\mu = \int f \, d\mu + \int g \, d\mu,$$

and

$$\int \alpha f \, d\mu = \alpha \int f \, d\mu \, .$$

Exercise 3.2.4. *For $f : \Omega \to [0, \infty]$ measurable, define*

$$s_n = n \mathbf{1}_{[f \ge n]} + \sum_{i=0}^{n2^n - 1} 2^{-n} i \mathbf{1}_{[2^{-n} i \le f < 2^{-n}(i+1)]}, \; n \ge 1 \, .$$

Show that $0 \le s_n \uparrow f$.

Proof of Theorem 3.2.6. For the first claim, let us first prove it for the case when f, g are simple. If

$$f = \sum_{i=1}^m \alpha_i \mathbf{1}_{A_i} \text{ and } g = \sum_{j=1}^n \beta_j \mathbf{1}_{B_j},$$

where $A_1, \ldots, A_m \in \mathcal{A}$ are disjoint, $A_1 \cup \ldots \cup A_m = \Omega$ and likewise for $B_1, \ldots, B_n$, then

$$f + g = \sum_{i=1}^m \sum_{j=1}^n (\alpha_i + \beta_j) \mathbf{1}_{A_i \cap B_j} \, .$$

Theorem 3.2.1 shows

$$\int (f+g) \, d\mu = \sum_{i=1}^m \sum_{j=1}^n (\alpha_i + \beta_j) \mu(A_i \cap B_j) = \int f \, d\mu + \int g \, d\mu \, .$$

For general measurable f, g, Exc 3.2.4 and MCT prove the claim. The second claim follows in a similar way. □

Corollary 3.2.1. *If*

$$f = \sum_{i=1}^\infty \alpha_i \mathbf{1}_{A_i}$$

for $A_1, A_2, \ldots \in \mathcal{A}$ which are not necessarily disjoint and $\alpha_1, \alpha_2, \ldots \in [0, \infty]$, then

$$\int f \, d\mu = \sum_{i=1}^\infty \alpha_i \mu(A_i) \, .$$

Exercise 3.2.5. *Suppose $f : \Omega \to [0, \infty]$ is measurable with respect to a σ-field $\mathcal{A}_0 \subset \mathcal{A}$. Show that*

$$\int_{(\Omega,\mathcal{A}_0)} f \, d\mu = \int_{(\Omega,\mathcal{A})} f \, d\mu \,.$$

Exercise 3.2.6. *Suppose $f \geq 0$ is measurable.*

1. *If $\mu([f > 0]) > 0$, show that*

$$\int f \, d\mu > 0 \,.$$

2. *If $\mu([f = \infty]) > 0$, show that*

$$\int f \, d\mu = \infty \,.$$

3.3 Integration of measurable functions

For all $x \in \overline{\mathbb{R}}$, define

$$x^+ = x \vee 0 \text{ and } x^- = (-x) \vee 0 \,;$$

x^+ and x^- are the "positive" and "negative" parts of x, respectively. The negative part of a number should not be confused with the negative of a number. In fact,

$$x^+ \geq 0 \text{ and } x^- \geq 0 \text{ for all } x \in \overline{\mathbb{R}} \,.$$

Exercise 3.3.1. *Show that for all $x \in \overline{\mathbb{R}}$,*

$$x = x^+ - x^- \,.$$

Inspired by the above exercise, the integral of a function is defined as the difference of the integrals of its positive and negative parts, as follows.

Definition 3.3.1. *For a measurable $f : \Omega \to \overline{\mathbb{R}}$, the integral of f with respect to μ, denoted by $\int f \, d\mu$, is defined by*

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

whenever either $\int f^+ \, d\mu$ or $\int f^- \, d\mu$ is finite. If both $\int f^+ \, d\mu$ and $\int f^- \, d\mu$ are finite, then f is integrable.

The following notations will also be used to denote the integral of f with respect to μ:

$$\int_\Omega f\,d\mu,\ \int f(x)\,d\mu(x),\ \int f(x)\,\mu(dx) \text{ etc.} \tag{3.3.1}$$

For $A \in \mathcal{A}$, the "integral over A" for a measurable f means the following:

$$\int_A f\,d\mu = \int f\mathbf{1}_A\,d\mu, \tag{3.3.2}$$

whenever the integral on the right hand side is defined. Taking $A = \Omega$, the right hand side of (3.3.2) becomes simply the integral of f, which is consistent with the notation $\int_\Omega f\,d\mu$ in (3.3.1).

Exercise 3.3.2. *1. Show that the following are equivalent for a measurable f:*

  (a) f is integrable,
  (b) $|f|$ is integrable,
  (c) f^+ and f^- are integrable,
  (d) $\int |f|\,d\mu < \infty$,
  (e) the integral of f is defined and finite.

2. Show that 'f is integrable' and 'the integral of f is defined' are not the same. Though the former implies the latter, the converse is false.

Definition 3.3.2. *If $A \in \mathcal{A}$ is such that $\mu(A^c) = 0$, then anything which holds on A is said to hold* almost everywhere *or a.e.*

Theorem 3.3.1. *The following hold for f, g whose integrals are defined.*

1. *If $f \le g$ a.e., then*
$$\int f\,d\mu \le \int g\,d\mu\,.$$
2. *If $f = g$ a.e., then*
$$\int f\,d\mu = \int g\,d\mu\,.$$
3. *For $\alpha \in \mathbb{R}$,*
$$\int \alpha f\,d\mu = \alpha \int f\,d\mu\,.$$
4. *It holds that*
$$\left|\int f\,d\mu\right| \le \int |f|\,d\mu\,.$$

5. *If f, g are integrable, then $f + g$ is defined a.e. and is integrable, and*

$$\int (f+g)\, d\mu = \int f\, d\mu + \int g\, d\mu\,.$$

Proof. 1. If $f \le g$ a.e., then $f^+ \le g^+$ a.e. In the notation (3.3.2), it follows that

$$\int_{[f^+ > g^+]} f^+\, dP = 0\,.$$

Thus

$$\begin{aligned} \int f^+\, dP &= \int_{[f^+ \le g^+]} f^+\, dP \\ \big(\text{because } f^+\mathbf{1}_{[f^+ \le g^+]} \le g^+\mathbf{1}_{[f^+ \le g^+]} \le g^+\big) &\le \int g^+\, dP\,. \end{aligned}$$

A similar argument shows

$$\int f^-\, d\mu \ge \int g^-\, d\mu,$$

from which the claim follows.

2. Follows from 1.

3. Follows from the observation

$$(\alpha f)^+ = \alpha f^+ \text{ and } (\alpha f)^- = \alpha f^- \text{ if } \alpha \ge 0,$$

and

$$(\alpha f)^+ = -\alpha f^- \text{ and } (\alpha f)^- = -\alpha f^+ \text{ if } \alpha < 0\,.$$

4. The definition of integral implies

$$\begin{aligned} \left|\int f\, d\mu\right| &= \left|\int f^+\, d\mu - \int f^-\, d\mu\right| \\ &\le \int f^+\, d\mu + \int f^-\, d\mu \\ &= \int |f|\, d\mu, \end{aligned}$$

the last line following from Theorem 3.2.6.

5. Exc 3.3.2 implies that

$$\int |f|\, d\mu < \infty \text{ and } \int |g|\, d\mu < \infty\,.$$

Exc 3.2.6 shows f and g are finite a.e. Hence $f+g$ is defined a.e. The above along with the inequality $|f+g| \le |f| + |g|$ shows $f+g$ is integrable by Exc 3.3.2.

Since,

$$(f+g)^+ - (f+g)^- = f + g = f^+ - f^- + g^+ - g^-,$$

arranging terms we get

$$(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+ = h \text{ (say)}.$$

Theorem 3.2.6 shows

$$\begin{aligned} &\int (f+g)^+ \, d\mu + \int f^- \, d\mu + \int g^- \, d\mu && (3.3.3)\\ &= \int h \, d\mu \\ &= \int (f+g)^- \, d\mu + \int f^+ \, d\mu + \int g^+ \, d\mu \, . && (3.3.4) \end{aligned}$$

Since f, g and $f+g$ are integrable, Exc 3.3.2 implies the integrals of $f^\pm, g^\pm, (f+g)^\pm$ are all finite. Thus terms between (3.3.3) and (3.3.4) can be arranged, from which it follows that

$$\begin{aligned} &\int (f+g)^+ \, d\mu - \int (f+g)^- \, d\mu \\ &= \int f^+ \, d\mu - \int f^- \, d\mu + \int g^+ \, d\mu - \int g^- \, d\mu, \end{aligned}$$

which is the same as the claimed equality

$$\int (f+g) \, d\mu = \int f \, d\mu + \int g \, d\mu \, .$$

This completes the proof.

□

Remark 3.3.1. *As is the case for $f+g$ in 5. of the above theorem, it suffices for a function to be defined a.e. On the set where it is not defined, the function may be redefined as zero, for example, which doesn't really matter as long as that set has zero measure.*

Theorem 3.3.2 (Fatou's lemma)**.** *For measurable $f_1, f_2, \ldots \ge 0$,*

$$\int \left(\liminf_{n\to\infty} f_n \right) \, d\mu \le \liminf_{n\to\infty} \int f_n \, d\mu \, .$$

Proof. Let

$$g_n = \inf_{k \geq n} f_k,\ n = 1, 2, \ldots,$$

and

$$g_\infty = \liminf_{n \to \infty} f_n\,.$$

Then $0 \leq g_n \uparrow g_\infty$. MCT implies

$$\begin{aligned}\int g_\infty \, d\mu &= \lim_{n \to \infty} \int g_n \, d\mu \\ &= \liminf_{n \to \infty} \int g_n \, d\mu \\ &\leq \liminf_{n \to \infty} \int f_n \, d\mu,\end{aligned}$$

the last line following from the obvious observation that $g_n \leq f_n$. This completes the proof. ☐

Theorem 3.3.3 (Dominated convergence theorem)**.** *If $f_n \to f$ and $|f_n| \leq g$, where f_n, f, g are measurable and g is integrable, then f_n and f are integrable, and*

$$\lim_{n \to \infty} \int f_n \, d\mu = \int f \, d\mu\,.$$

Proof. The assumptions imply $f_n + g \geq 0$. Further

$$\begin{aligned}\int (f+g) \, d\mu &= \int \liminf_{n \to \infty} (f_n + g) \, d\mu \\ \text{(Fatou's lemma)} &\leq \liminf_{n \to \infty} \int (f_n + g) \, d\mu \\ &= \int g \, d\mu + \liminf_{n \to \infty} \int f_n \, d\mu\,.\end{aligned}$$

Subtracting $\int g \, d\mu$ from both sides, which is finite, we get

$$\int f \, d\mu \leq \liminf_{n \to \infty} \int f_n \, d\mu\,. \tag{3.3.5}$$

A similar argument with Fatou's lemma shows

$$\int (g-f) \, d\mu \leq \liminf_{n \to \infty} \int (g - f_n) d\,\mu = \int g \, d\mu - \limsup_{n \to \infty} \int f_n \, d\mu,$$

which implies

$$\int f \, d\mu \geq \limsup_{n \to \infty} \int f_n \, d\mu\,.$$

The proof follows by combining the above with (3.3.5). ☐

Remark 3.3.2. *The trio of MCT, Fatou's lemma and DCT is ubiquitous in mathematics.*

3.4 Inequalities

As before, $(\Omega, \mathcal{A}, \mu)$ is a measure space, and all functions talked about are measurable functions from Ω to $\overline{\mathbb{R}}$.

Theorem 3.4.1 (Hölder). *For* $p, q > 1$ *such that* $\frac{1}{p} + \frac{1}{q} = 1$,

$$\int |fg| \, d\mu \le \left(\int |f|^p \, d\mu \right)^{1/p} \left(\int |g|^q \, d\mu \right)^{1/q} .$$

The proof uses the following lemma. Unless mentioned otherwise, $f'(x)$ denotes the derivative of f at x if it exists, with the understanding that one may have to look at the right or the left derivative at the endpoints of the interval on which the function is defined. This standard convention is adopted throughout.

Lemma 3.4.1. *For* $0 < \lambda < 1$, $a, b \ge 0$ *and* $t > 0$,

$$a^\lambda b^{1-\lambda} \le \lambda t^{\lambda-1} a + (1-\lambda) t^\lambda b \, .$$

Proof. Assume without loss of generality that $a, b > 0$ for else the left hand side vanishes. Define

$$f(x) = \lambda x^{\lambda-1} a + (1-\lambda) x^\lambda b \, , x > 0 \, .$$

Differentiating with respect to x, we get

$$\begin{aligned} f'(x) &= \lambda(\lambda - 1) x^{\lambda-2} a + (1-\lambda)\lambda x^{\lambda-1} b \\ &= \lambda(1-\lambda) x^{\lambda-2} (bx - a) \, . \end{aligned}$$

Thus $f' \ge 0$ on $[a/b, \infty)$ and $f' \le 0$ on $(0, a/b]$. Hence for any $t > 0$,

$$f(t) \ge f(a/b) = \lambda (a/b)^{\lambda-1} a + (1-\lambda)(a/b)^\lambda b = a^\lambda b^{1-\lambda},$$

which completes the proof. ☐

Proof of Theorem 3.4.1. Fix measurable functions f, g and $p, q > 1$ such that

$$\frac{1}{p} + \frac{1}{q} = 1 \, . \tag{3.4.1}$$

Assume without loss of generality that

$$\int |f|^p \, d\mu > 0 \text{ and } \int |g|^q \, d\mu > 0,$$

because otherwise either $f = 0$ a.e. or $g = 0$ a.e. Next assume without loss of generality that the above quantities are finite as well, because otherwise the right hand side of the claimed inequality is ∞.

Use Lemma 3.4.1 with $\lambda = \frac{1}{p}$, $a = |f|^p$ and $b = |g|^q$ to get for a fixed $t > 0$

$$|f|\,|g| \le \frac{1}{p} t^{-1/q} |f|^p + \frac{1}{q} t^{1/p} |g|^q \,.$$

Thus,

$$\int |fg|\,d\mu \le \frac{1}{p} t^{-1/q} \int |f|^p\,d\mu + \frac{1}{q} t^{1/p} \int |g|^q\,d\mu \,.$$

The above holds for all $t > 0$. Putting

$$t = \left(\int |g|^q\,d\mu \right)^{-1} \int |f|^p\,d\mu,$$

the right hand side becomes

$$\frac{1}{p} \left(\int |f|^p\,d\mu \right)^{1-1/q} \left(\int |g|^q\,d\mu \right)^{1/q} + \frac{1}{q} \left(\int |f|^p \right)^{1/p} \left(\int |g|^q\,d\mu \right)^{1-1/p} .$$

Since (3.4.1) implies the above is the same as the right hand side of the claimed inequality, the proof follows. ☐

Theorem 3.4.2 (Cauchy-Schwarz)**.** *For f, g measurable,*

$$\int |fg|\,d\mu \le \left(\int f^2\,d\mu \int g^2\,d\mu \right)^{1/2} .$$

Proof. Follows from Theorem 3.4.1 by putting $p = q = 2$. ☐

Theorem 3.4.3 (Minkowski)**.** *For $1 \le p < \infty$,*

$$\left(\int |f+g|^p\,d\mu \right)^{1/p} \le \left(\int |f|^p\,d\mu \right)^{1/p} + \left(\int |g|^p\,d\mu \right)^{1/p} .$$

Proof. Without loss of generality, assume that the right hand side of the claimed inequality is finite. Then

$$\int |f+g|^p\,d\mu \le \int (|f|+|g|)^p\,d\mu \le 2^p \int (|f|^p + |g|^p)\,d\mu < \infty \,.$$

Denote

$$c = \left(\int |f+g|^p\,d\mu \right)^{1/p} < \infty \,.$$

Assume without loss of generality that $c > 0$ for else the claimed inequality is trivial. In addition, assume without loss of generality that $p > 1$ because for $p = 1$ the claim follows trivially from

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

Write

$$\begin{aligned}
c^p &= \int |f+g|^p \, d\mu \\
&= \int |f+g||f+g|^{p-1} \, d\mu \\
&\leq \int |f||f+g|^{p-1} \, d\mu + \int |g||f+g|^{p-1} \, d\mu \\
\text{(Hölder inequality)} \quad &\leq \left(\int |f|^p \, d\mu \right)^{1/p} \left(\int |f+g|^{(p-1)q} \, d\mu \right)^{1/q} \\
&\quad + \left(\int |g|^p \, d\mu \right)^{1/p} \left(\int |f+g|^{(p-1)q} \, d\mu \right)^{1/q} \\
&= c^{p/q} \left[\left(\int |f|^p \, d\mu \right)^{1/p} + \left(\int |g|^p \, d\mu \right)^{1/p} \right],
\end{aligned}$$

where $q = (1 - 1/p)^{-1}$. Therefore

$$\left(\int |f|^p \, d\mu \right)^{1/p} + \left(\int |g|^p \, d\mu \right)^{1/p} \geq c^{p-p/q} = c,$$

which completes the proof. □

Once again, the notation (3.3.2) is used in the following theorem.

Theorem 3.4.4 (Markov)**.** *For $f \geq 0$ and $a \in (0, \infty)$,*

$$\mu\left([f \geq a]\right) \leq \frac{1}{a} \int f \, d\mu \, .$$

Proof. By writing

$$\begin{aligned}
\int f \, d\mu &\geq \int_{[f \geq a]} f \, d\mu \\
&\geq \int_{[f \geq a]} a \, d\mu \\
&= a \, \mu\left([f \geq a]\right),
\end{aligned}$$

the proof follows. □

Remark 3.4.1. *Theorem 3.4.4 is also known as Chebyshev's inequality in several references.*

Definition 3.4.1. *For $-\infty \leq a < b \leq \infty$, a function $\varphi : (a, b) \to \mathbb{R}$ is* convex *if*

$$\varphi\left(\lambda x + (1-\lambda) y\right) \leq \lambda \varphi(x) + (1-\lambda)\varphi(y) \text{ for all } x, y \in (a, b) \, .$$

Exercise 3.4.1. *If* $(a,b) \subset \mathbb{R}$ *and* $\varphi : (a,b) \to \mathbb{R}$ *is convex, show that*

1. φ *is continuous and hence Borel,*
2. *for all* $x \in (a,b)$*,*
$$\varphi'_+(x) = \lim_{h \downarrow 0} \frac{\varphi(x+h) - \varphi(x)}{h}$$
exists,
3. *and for all* $x, x_0 \in (a,b)$*,*
$$\varphi(x) \geq \varphi(x_0) + (x - x_0)\varphi'_+(x_0)\,. \tag{3.4.2}$$

Theorem 3.4.5 (Jensen)**.** *Suppose* $(\Omega, \mathcal{A}, P)$ *is a probability space,* $f : \Omega \to (a,b)$ *is measurable for some* $-\infty \leq a < b \leq \infty$ *and* $\varphi : (a,b) \to \mathbb{R}$ *is convex. If* f *and* $\varphi \circ f$ *are integrable functions, then*

$$a < \int f\,dP < b, \tag{3.4.3}$$

and

$$\int \varphi(f)\,dP \geq \varphi\left(\int f\,dP\right)\,.$$

Proof. Since $b > f$ and $P(\Omega) > 0$, Exc 3.2.6 shows that

$$0 < \int (b - f)\,dP = b - \int f\,dP,$$

the equality following from $P(\Omega) = 1$. As $\int f\,dP$ is a finite quantity, it can be taken to the left hand side, which shows

$$\int f\,dP < b\,.$$

A similar argument proves

$$\int f\,d\mu > a,$$

(3.4.3) follows from which.

Letting $x_0 = \int f\,dP$, (3.4.2) shows

$$\varphi(f) \geq \varphi(x_0) + (f - x_0)\varphi'_+(x_0)\,.$$

Thus

$$\begin{aligned}\int \varphi(f)\,dP &\geq \int \big(\varphi(x_0) + (f - x_0)\varphi'_+(x_0)\big)\,dP \\ (\text{since } P(\Omega) = 1) &= \varphi(x_0) + \varphi'_+(x_0)\left(\int f\,dP - x_0\right) \\ &= \varphi(x_0),\end{aligned}$$

and hence the proof. □

3.5 The L^p space

As usual, $(\Omega, \mathcal{A}, \mu)$ is a measure space. All functions talked about are measurable functions from Ω to $\mathbb{R}$, unless mentioned otherwise.

Definition 3.5.1. *For $1 \le p < \infty$, define*

$$L^p(\Omega, \mathcal{A}, \mu) = \left\{ f : \Omega \to \mathbb{R} \text{ measurable } : \int |f|^p \, d\mu < \infty \right\},$$

and for all $f \in L^p(\Omega, \mathcal{A}, \mu)$, define

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

For $f, f_1, f_2, \ldots \in L^p(\Omega, \mathcal{A}, \mu)$, we say $f_n \to f$ in L^p if

$$\|f_n - f\|_p \to 0 \text{ as } n \to \infty .$$

The mention of $\mathcal{A}$ or μ in $L^p(\Omega, \mathcal{A}, \mu)$ will be suppressed whenever the same is obvious from the context. For $f, g \in L^p(\Omega)$, f and g will be identified with each other if $f = g$ a.e. In other words, $L^p(\Omega)$ will refer to the set of partitions induced by the equivalence relation $\sim$ which is defined as follows: for $f, g \in L^p(\Omega)$,

$$f \sim g \iff f = g \text{ a.e.}$$

Let us fix $1 \le p < \infty$ for this subsection.

Theorem 3.5.1. *The collection $L^p(\Omega)$ is a vector space over $\mathbb{R}$ and $\| \cdot \|_p$ is a norm on that.*

Proof. The only non-trivial property of $\| \cdot \|_p$ to be checked is the triangle inequality, which follows from Theorem 3.4.3. ☐

Corollary 3.5.1. *If*

$$d(f, g) = \|f - g\|_p, \; f, g \in L^p(\Omega),$$

then d is a metric on $L^p(\Omega)$, in which, two functions are identified if they are equal a.e.

Definition 3.5.2. *For measurable functions $f_1, f_2, \ldots, f_\infty$ from Ω to $\overline{\mathbb{R}}$, we say $f_n \to f_\infty$ a.e. if*

$$\mu \left(\left\{ \omega \in \Omega : \lim_{n \to \infty} f_n(\omega) = f_\infty(\omega) \right\}^c \right) = 0 .$$

Exercise 3.5.1. *If $f_n \to f$ a.e. and $|f_n| \le g$, where f_n, f, g are measurable and g is integrable, then show that*

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

and

$$f_n \to f \text{ in } L^1 .$$

Theorem 3.5.2 (Scheffé's lemma). *Suppose $f, f_1, f_2, \ldots$ are measurable functions such that $0 \le f_n \to f$ a.e. Then $f_n \to f$ in L^1 if and only if*

$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu < \infty .$$

Proof. For the 'if' part, write

$$\|f_n - f\|_1 = \int f_n \, d\mu + \int f \, d\mu - 2 \int (f_n \wedge f) \, d\mu . \tag{3.5.1}$$

Exc 3.5.1 along with the fact $0 \le f_n \wedge f \le f$ implies

$$\lim_{n\to\infty} \int (f_n \wedge f) \, d\mu = \int f \, d\mu .$$

This along with the assumed hypothesis implies that the right hand side of (3.5.1) goes to zero as $n \to \infty$. Hence the 'if' part follows.

The 'only if' part follows trivially from

$$\left| \int f_n \, d\mu - \int f \, d\mu \right| = \big| \|f_n\|_1 - \|f\|_1 \big| \le \|f_n - f\|_1,$$

the inequality following from Theorem 3.5.1 with $p = 1$. This completes the proof. ☐

Definition 3.5.3. *For sets $A_1, A_2, A_3, \ldots$, define*

$$\limsup_n A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k,$$

and

$$\liminf_n A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k .$$

Theorem 3.5.3 (Borel-Cantelli lemma). *Suppose that $\sum_{n=1}^{\infty} \mu(A_n) < \infty$ for $A_1, A_2, \ldots \in \mathcal{A}$. Then*

$$\mu\left(\limsup_n A_n \right) = 0 .$$

Proof. Let $B = \limsup A_n$ and

$$B_n = \bigcup_{k=n}^{\infty} A_k,\ n \geq 1\,.$$

Then $B_n \downarrow B$. Since

$$\mu(B_1) = \mu\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} \mu(A_k) < \infty,$$

it follows that $\mu(B_n) \downarrow \mu(B)$. Hence it suffices to show $\mu(B_n) \downarrow 0$.

Observing that for all $n = 1, 2, \ldots,$

$$\mu(B_n) \leq \sum_{k=n}^{\infty} \mu(A_k),$$

and that the right hand side goes to zero as $n \to \infty$, the proof follows. □

Theorem 3.5.4. *For measurable $f, f_1, f_2, \ldots$ from Ω to $\mathbb{R}$, $f_n \to f$ a.e. if and only if*

$$\mu\left(\limsup_n \left[|f_n - f| > \varepsilon\right]\right) = 0 \textit{ for all } \varepsilon > 0\,. \tag{3.5.2}$$

Proof. Define

$$g = \limsup_{n\to\infty} |f_n - f|\,.$$

Then $[f_n \to f] = [g = 0]$. For all $\varepsilon > 0$,

$$\begin{aligned}[g > \varepsilon] \subset [|f_n - f| > \varepsilon \text{ for infinitely many } n] &= \limsup_n \left[|f_n - f| > \varepsilon\right] \qquad (3.5.3)\\ &\subset [g \geq \varepsilon]\,. \qquad (3.5.4)\end{aligned}$$

Thus if $g = 0$ a.e., then (3.5.4) implies

$$\mu\left(\limsup_n \left[|f_n - f| > \varepsilon\right]\right) \leq \mu([g \geq \varepsilon]) = 0,$$

which proves the ‘only if’ part.

For the ‘if’ part, assume (3.5.2). For all $\varepsilon > 0$, (3.5.3) implies

$$\mu\left([g > \varepsilon]\right) = 0\,.$$

Since $[g > 1/k] \uparrow [g > 0]$ as $k \to \infty$, it thus follows that $g = 0$ a.e. Hence the proof follows. □

The following is a trivial consequence of Theorems 3.5.3 and 3.5.4.

Corollary 3.5.2. *For measurable functions $f_1, f_2, \ldots, f_\infty$ from Ω to $\mathbb{R}$ satisfying*

$$\sum_{n=1}^{\infty} \mu\left([|f_n - f_\infty| > \varepsilon]\right) < \infty \text{ for all } \varepsilon > 0,$$

it holds that $f_n \to f_\infty$ a.e.

Theorem 3.5.5. *If $f_n \to f$ in L^p, then $\{f_n\}$ has a subsequence which converges to f a.e.*

Proof. The hypothesis implies the existence of $1 \le n_1 < n_2 < \ldots$ such that

$$\|f_{n_k} - f\|_p \le 4^{-k},\ k \ge 1\,.$$

Then for $\varepsilon > 0$,

$$\begin{aligned} \mu\left([|f_{n_k} - f| > \varepsilon]\right) &= \mu\left([|f_{n_k} - f|^p > \varepsilon^p]\right) \\ \text{(Markov inequality)} &\le \varepsilon^{-p} \int |f_{n_k} - f|^p \, d\mu \\ &= \varepsilon^{-p} \|f_{n_k} - f\|_p^p \\ &\le \varepsilon^{-p} 4^{-kp}\,. \end{aligned}$$

Hence

$$\sum_{k=1}^{\infty} \mu\left([|f_{n_k} - f| > \varepsilon]\right) < \infty \text{ for all } \varepsilon > 0\,.$$

Corollary 3.5.2 completes the proof. $\square$

Example 3.5.1. *Let $\Omega = (0,1]$, $\mathcal{A} = \mathcal{B}((0,1])$ and μ be the restriction of Lebesgue measure to $(0,1]$. Define for all $\omega \in \Omega$,*

$$\begin{aligned} f_1(\omega) &= \mathbf{1}\left(0 < \omega \le \frac{1}{2}\right), \\ f_2(\omega) &= \mathbf{1}\left(\frac{1}{2} < \omega \le 1\right), \\ f_3(\omega) &= \mathbf{1}\left(0 < \omega \le \frac{1}{4}\right), \\ f_4(\omega) &= \mathbf{1}\left(\frac{1}{4} < \omega \le \frac{1}{2}\right), \\ f_5(\omega) &= \mathbf{1}\left(\frac{1}{2} < \omega \le \frac{3}{4}\right), \\ f_6(\omega) &= \mathbf{1}\left(\frac{3}{4} < \omega \le 1\right), \\ &\vdots \end{aligned}$$

and so on. Show that $f_n \to 0$ *in* L^p *for all* $1 \le p < \infty$ *and*

$$\liminf_{n\to\infty} f_n = 0 < 1 = \limsup_{n\to\infty} f_n \text{ a.e.}$$

Exercise 3.5.2. *Show that for a fixed* $1 \le p < \infty$*, neither of convergence in* L^p *and a.e. convergence implies the other.*

Exercise 3.5.3. *Suppose* $f_n \to g$ *in* L^p *and* $f_n \to h$ *a.e. Show that* g *and* h *are the same, that is,*

$$g = h \text{ a.e.}$$

Exercise 3.5.4. *Show by citing a counter-example that a.e. convergence is not metrizable, that is, there does not exist a metric* d *on the space of all measurable functions from* Ω *to* $\mathbb{R}$ *such that convergence in* d *is equivalent to a.e. convergence.*
Hint. The following steps may be executed.
Step 1. *Show that for* $\alpha_n \in \mathbb{R}$*,* $\alpha_n \to 0$ *if and only if every subsequence of* α_n *has a further subsequence which converges to* 0*.*
Step 2. *Use Step 1 to argue that in a metric space* (Σ, d)*, for* $x, x_1, x_2, \ldots \ldots \in \Sigma$*,* $x_n \to x$ *if and only if every subsequence of* x_n *has a further subsequence converging to* x*.*
Step 3. *Consider* f_n *as in Example 3.5.1. Show that every subsequence of* f_n *converges to the zero function in* L^1*.*
Step 4. *Show (using Theorem 3.5.5) that every subsequence of* f_n *has a further subsequence which converges to* 0 *a.e. However,* f_n *does not converge to* 0 *a.e.*
Step 5. *Use Steps 2 and 4 to argue that a.e. convergence on this* $(\Omega, \mathcal{A}, \mu)$ *is not metrizable.*

Theorem 3.5.6. *The metric space* $L^p(\Omega)$ *is complete.*

Proof. Let $\{f_n\}$ be a Cauchy sequence in $L^p(\Omega)$, that is, for all $\varepsilon > 0$, there exists N such that

$$\|f_m - f_n\|_p \le \varepsilon,\ m, n \ge N\,.$$

This allows us to choose $1 \le n_1 < n_2 < \ldots$ such that

$$\|f_i - f_j\|_p \le 4^{-k} \text{ for all } i, j \ge n_k\,. \tag{3.5.5}$$

This by an appeal to Minkowski implies

$$\sup_{n\ge1} \|f_n\|_p \le \frac{1}{4} + \max_{1\le i\le n_1} \|f_i\|_p < \infty\,. \tag{3.5.6}$$

Another immediate consequence of (3.5.5) is that

$$\left\|f_{n_{k+1}} - f_{n_k}\right\|_p \le 4^{-k},\ k \ge 1\,.$$

Arguments similar to those in the proof of Theorem 3.5.5, using the above and the Markov inequality, imply that

$$\mu\left(\left[|f_{n_{k+1}} - f_{n_k}| > 2^{-k}\right]\right) \le 2^{-kp},\ k \ge 1\,.$$

Thus

$$\sum_{k=1}^{\infty} \mu\left(\left[|f_{n_{k+1}} - f_{n_k}| > 2^{-k}\right]\right) < \infty\,.$$

Theorem 3.5.3 implies for a.e. $\omega \in \Omega$,

$|f_{n_{k+1}}(\omega) - f_{n_k}(\omega)| \le 2^{-k}$ for all but finitely many k's, depending upon ω,

and therefore,

$$\sum_{k=1}^{\infty} |f_{n_{k+1}} - f_{n_k}| < \infty \text{ a.e.}$$

Letting

$$\Omega_0 = \left[\sum_{k=1}^{\infty} |f_{n_{k+1}} - f_{n_k}| < \infty\right],$$

it follows that for all $\omega \in \Omega_0$, $\{f_{n_k}(\omega) : k \ge 1\}$ is a Cauchy sequence in $\mathbb{R}$. Since $\mathbb{R}$ is complete, $\lim_{k\to\infty} f_{n_k}(\omega)$ exists for all $\omega \in \Omega_0$.

Let

$$f = \limsup_{k\to\infty} f_{n_k}\,.$$

Since $f_{n_k} \to f$ a.e.,

$$\begin{aligned} \int |f|^p\,d\mu &= \int \liminf_{k\to\infty} |f_{n_k}|^p\,d\mu \\ \text{(Fatou's lemma)} &\le \liminf_{k\to\infty} \int |f_{n_k}|^p\,d\mu \\ &< \infty, \end{aligned}$$

(3.5.6) implying the last line. Thus $f \in L^p(\Omega)$.

For a fixed $k = 1, 2, 3, \ldots$, (3.5.5) implies

$$\|f_{n_k} - f_{n_l}\|_p \le 4^{-k},\ l \ge k\,.$$

Thus

$$\begin{aligned} \|f_{n_k} - f\|_p^p &= \int |f_{n_k} - f|^p\,d\mu \\ &= \int \liminf_{l\to\infty} |f_{n_k} - f_{n_l}|^p\,d\mu \\ \text{(Fatou's lemma)} &\le \liminf_{l\to\infty} \int |f_{n_k} - f_{n_l}|^p\,d\mu \\ &\le 4^{-kp}\,. \end{aligned}$$

For $n \geq n_k$, $\|f_n - f_{n_k}\|_p \leq 4^{-k}$ by (3.5.5); Minkowski implies for such n,

$$\|f_n - f\|_p \leq \|f_n - f_{n_k}\|_p + \|f_{n_k} - f\|_p \leq 2^{1-2k} .$$

This shows $f_n \to f$ in L^p. In other words, every Cauchy sequence in $L^p(\Omega)$ is convergent. This completes the proof. □

Definition 3.5.4. *For measurable $f : \Omega \to \overline{\mathbb{R}}$, define*

$$\|f\|_\infty = \inf\{0 \leq \alpha \leq \infty : |f| \leq \alpha \text{ a.e.}\} .$$

As before, $L^\infty(\Omega) = \{f : \|f\|_\infty < \infty\}$ with the understanding that two functions are considered the same if they are equal a.e. For $f, f_1, f_2, \ldots \in L^\infty(\Omega)$, $f_n \to f$ in L^∞ if $\|f_n - f\|_\infty \to 0$.

Theorem 3.5.7. *For a measurable $f : \Omega \to \overline{\mathbb{R}}$, $|f| \leq \|f\|_\infty$ a.e.*

Proof. The definition of the L^∞-norm implies there exists $\alpha_n \downarrow \|f\|_\infty$ such that for all $n = 1, 2, \ldots$,

$$|f| \leq \alpha_n \text{ a.e.}$$

Since $\alpha_n \downarrow \|f\|_\infty$, we have

$$[|f| > \alpha_n] \uparrow [|f| > \|f\|_\infty] .$$

Continuity from below implies

$$\mu\left(|f| > \|f\|_\infty\right) = \lim_{n\to\infty} \mu\left(|f| > \alpha_n\right) = 0 .$$

Hence the proof. □

Exercise 3.5.5. *Show that $L^\infty(\Omega)$ is a complete metric space.*

Exercise 3.5.6. *Show that Hölder's inequality holds with $p = 1$ and $q = \infty$, that is,*

$$\|fg\|_1 \leq \|f\|_1 \|g\|_\infty .$$

In other words, prove that

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

where $1 \leq p, q \leq \infty$ are such that

$$\frac{1}{p} + \frac{1}{q} = 1,$$

$\frac{1}{\infty}$ being interpreted as 0 above.

Exercise 3.5.7. *If $\mu(\Omega) < \infty$, show that*

$$\lim_{p\to\infty} \|f\|_p = \|f\|_\infty .$$

Exercise 3.5.8. *Show the following for $f_1, \ldots, f_\infty$.*

1. *If $f_n \to f_\infty$ in L^∞, then show that there exist measurable functions $g_1, \ldots, g_\infty$ such that*

$$\lim_{n\to\infty} \sup_{\omega\in\Omega} |g_n(\omega) - g_\infty(\omega)| = 0,$$

and

$$g_n = f_n \text{ a.e. for } n = 1, 2, \ldots, \infty\,.$$

Since f_n and g_n are considered identical elements of $L^\infty(\Omega)$, convergence in L^∞ essentially means uniform convergence.

2. *If $f_n \to f_\infty$ a.e., then show that there exist measurable functions $g_1, \ldots, g_\infty$ such that*

$$\lim_{n\to\infty} g_n(\omega) = g_\infty(\omega) \text{ for all } \omega \in \Omega,$$

and

$$g_n = f_n \text{ a.e. for } n = 1, 2, \ldots, \infty\,.$$

In other words, a.e. convergence essentially means pointwise convergence.

3. *If $f_n \to f_\infty$ in L^∞, show that $f_n \to f_\infty$ a.e.*

Exercise 3.5.9.

1. *For $1 \le p < q \le \infty$, show that neither of $L^p(\mathbb{R}, \lambda)$ and $L^q(\mathbb{R}, \lambda)$ is a subset of the other, where λ is the Lebesgue measure.*

2. *If $\mu(\Omega) < \infty$, show that*

$$L^p(\Omega) \supset L^q(\Omega) \text{ if } 1 \le p \le q \le \infty\,.$$

3. *If $\mu(\Omega) = 1$, then show that*

$$\|f\|_p \le \|f\|_q \text{ if } 1 \le p \le q \le \infty\,.$$

This is known as Lyapunov's inequality.

4. *Let $\ell^p = L^p\big(\mathbb{N}, 2^{\mathbb{N}}, \mu\big)$ where μ is the counting measure, for $1 \le p \le \infty$. In other words,*

$$\ell^p = \left\{(x_1, x_2, x_3, \ldots) \in \mathbb{R}^{\mathbb{N}} : \sum_{n=1}^{\infty} |x_n|^p < \infty\right\} \text{ for } 1 \le p < \infty$$

and

$$\ell^\infty = \left\{(x_1, x_2, x_3, \ldots) \in \mathbb{R}^{\mathbb{N}} : \sup_{n\ge 1} |x_n| < \infty\right\}.$$

Show that

$$\ell^p \subsetneq \ell^q \text{ if } 1 \le p < q \le \infty\,.$$

Exercise 3.5.10. *Suppose* $1 \le p \le q \le \infty$ *and* $f_1, f_2, \ldots \in L^p(\Omega) \cap L^q(\Omega)$. *If*

$$f_n \to g_1 \text{ in } L^p,$$
$$f_n \to g_2 \text{ in } L^q$$

and

$$f_n \to g_3 \text{ a.e.},$$

show that $g_1 = g_2 = g_3$ *a.e.*

3.6 Lebesgue integration

Let μ be the Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. As in (2.3.1), denote by μ^* the outer measure of μ where $\mathcal{F}$ therein is replaced by $\mathcal{B}(\mathbb{R})$, that is,

$$\mu^*(E) = \inf\{\mu(A) : E \subset A,\, A \in \mathcal{B}(\mathbb{R})\},\, E \in 2^{\mathbb{R}}.$$

As in (2.3.4), let

$$\mathcal{L}(\mathbb{R}) = \left\{A \in 2^{\mathbb{R}} : \mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \text{ for all } E \in 2^{\mathbb{R}}\right\}.$$

Lemma 2.3.5 shows that $\mathcal{L}(\mathbb{R})$ is a σ-field an μ^* is a measure on $(\mathbb{R}, \mathcal{L}(\mathbb{R}))$, which agrees with μ on $\mathcal{B}(\mathbb{R})$ by Lemma 2.3.1.

The following is a minor observation.

Theorem 3.6.1. *If* $A \in \mathcal{L}(\mathbb{R})$ *is such that* $\mu(A) = 0$, *then* $B \in \mathcal{L}(\mathbb{R})$ *for all* $B \subset A$.

Proof. Suppose $\mu(A) = 0$ for some $A \in \mathcal{L}(\mathbb{R})$ and $B \subset A$. Since μ^* is monotone by definition,

$$\mu^*(B) \le \mu^*(A) = 0.$$

Thus for any $E \in 2^{\mathbb{R}}$,

$$\mu^*(E \cap B) + \mu^*(E \cap B^c) \le \mu^*(B) + \mu^*(E \cap B^c) = \mu^*(E \cap B^c) \le \mu^*(E);$$

(2.3.5) shows that $B \in \mathcal{L}(\mathbb{R})$. This completes the proof. ☐

Denote

$$\lambda(A) = \mu^*(A),\, A \in \mathcal{L}(\mathbb{R}).$$

Thus $(\mathbb{R}, \mathcal{L}(\mathbb{R}), \lambda)$ is a measure space.

Lemma 3.6.1. *If* $A \in \mathcal{L}(\mathbb{R})$ *is such that* $\lambda(A) = 0$, *then for any function* $g : \mathbb{R} \to \mathbb{R}$, $g\mathbf{1}_A$ *is* $\mathcal{L}(\mathbb{R})$*-measurable.*

Proof. If $h = g\mathbf{1}_A$, then for any $E \subset \mathbb{R} \setminus \{0\}$, $h^{-1}E \subset A$ and hence by Theorem 3.6.1, $h^{-1}E \in \mathcal{L}(\mathbb{R})$. For $E \subset \mathbb{R}$ such that $0 \in E$,

$$h^{-1}E = \left(h^{-1}(E^c)\right)^c \in \mathcal{L}(\mathbb{R})$$

because $h^{-1}(E^c) \in \mathcal{L}(\mathbb{R})$. This completes the proof. ☐

Remark 3.6.1. *Lemma 3.6.1 holds for any function g, including which are not measurable.*

Definition 3.6.1. *For $[a,b] \subset \mathbb{R}$, $f : [a,b] \to \overline{\mathbb{R}}$ is* Lebesgue measurable *if it is measurable with respect to the σ-field $\mathcal{L}([a,b])$, defined by*

$$\mathcal{L}([a,b]) = \{A \in \mathcal{L}(\mathbb{R}) : A \subset [a,b]\},$$

that is,

$$f^{-1}B \in \mathcal{L}([a,b]) \text{ for all } B \in \mathcal{B}(\overline{\mathbb{R}})\,.$$

If in addition, $\int_{[a,b]} |f|\,d\lambda < \infty$, then f is Lebesgue integrable *and $\int_{[a,b]} f\,d\lambda$ is the* Lebesgue integral *of f on $[a,b]$. The notations*

$$\int_a^b f(x)\,dx,\ \int_a^b f\,d\lambda \text{ etc.}$$

are also used for denoting the Lebesgue integral. As usual, define

$$\int_A f(x)\,dx = \int f\mathbf{1}_A\,d\lambda,\ A \in \mathcal{B}([a,b])\,.$$

In particular, if $A = [\alpha,\beta] \subset [a,b]$, then there is no conflict of notation in defining

$$\int_\alpha^\beta f\,d\lambda = \int_A f\,d\lambda\,.$$

Let us fix $[a,b] \subset \mathbb{R}$ for this subsection. Henceforth, "integrable on $[a,b]$" will mean Lebesgue integrable.

Exercise 3.6.1. *Suppose $f : [a,b] \to \mathbb{R}$ is Lebesgue integrable. Define*

$$g(x) = \int_a^x f(t)\,dt,\ a \le x \le b\,.$$

1. *Show that g is a continuous function.*
2. *If f is continuous at x, show that g is differentiable at x and its derivative at x equals $f(x)$.*

Theorem 3.6.2. *If $f : [a,b] \to \mathbb{R}$ is bounded and Riemann integrable, then f is Lebesgue integrable on $[a,b]$ and its Lebesgue integral equals its Riemann integral.*

Proof. Let $[a,b] \subset \mathbb{R}$ and $f : [a,b] \to \mathbb{R}$ be bounded and Riemann integrable. Let R denote its Riemann integral. Then

$$\lim_{n\to\infty} (b-a)2^{-n} \sum_{i=1}^{2^n} \inf\{f(x) : a+2^{-n}(i-1)(b-a) < x \le a+2^{-n}i(b-a)\} = R \tag{3.6.1}$$

and

$$\lim_{n\to\infty} (b-a)2^{-n} \sum_{i=1}^{2^n} \sup\{f(x) : a+2^{-n}(i-1)(b-a) < x \le a+2^{-n}i(b-a)\} = R\,. \tag{3.6.2}$$

Defining

$$P_n = \big\{\{a\}, (a, a+2^{-n}(b-a)], \ldots, (b-2^{-n}(b-a), b]\big\}\,,$$

$$\begin{aligned} f_n^L &= \sum_{A\in P_n} \mathbf{1}_A \inf_{x\in A} f(x), \\ \text{and } f_n^U &= \sum_{A\in P_n} \mathbf{1}_A \sup_{x\in A} f(x), \end{aligned}$$

(3.6.1) and (3.6.2) become

$$\lim_{n\to\infty} \int_{[a,b]} f_n^L \, d\lambda = R = \lim_{n\to\infty} \int_{[a,b]} f_n^U \, d\lambda\,.$$

This in view of the observation

$$f_1^L \le f_2^L \le \ldots \le f \le \ldots \le f_2^U \le f_1^U, \tag{3.6.3}$$

show that $\{g_n\}$, defined by

$$g_1 = f_1^L,\ g_2 = f_1^U,\ g_3 = f_2^L,\ g_4 = f_2^U, \ldots,$$

is a Cauchy sequence in $L^1\left([a,b], \mathcal{B}([a,b]), \lambda\right)$.

Theorem 3.5.6 shows there exists $g \in L^1([a,b], \mathcal{B}([a,b]), \lambda)$ such that $g_n \to g$ in L^1. Theorem 3.5.5 shows that each of $\{g_{2n-1}\}_{n\ge1}$ and $\{g_{2n}\}_{n\ge1}$ has a subsequence that converges to g in L^1. In other words, there exist $1 \le m_1 < m_2 < \ldots$ such that

$$f_{m_k}^L \to g \text{ a.e.}$$

and $1 \le n_1 < n_2 < \ldots$ such that

$$f_{n_k}^U \to g \text{ a.e.}$$

Denoting $A = [f^L_{m_k} \to g] \cap [f^U_{n_k} \to g]$, (3.6.3) shows that $f = g$ on A. Further, $A \in \mathcal{B}([a,b])$ and $\lambda([a,b] \setminus A) = 0$. Thus

$$f = g\mathbf{1}_A + f\mathbf{1}_{[a,b]\setminus A};$$

Lemma 3.6.1 shows $f\mathbf{1}_{[a,b]\setminus A}$ is $\mathcal{L}([a,b])$-measurable. Thus so is f. Further, $f = g$ a.e. Hence $g_n \to f$ in L^1. This shows

$$\int_{[a,b]} f\, d\lambda = \lim_{n\to\infty} \int_{[a,b]} g_n\, d\lambda = R,$$

which completes the proof. $\square$

Lebesgue integration achieves the first goal mentioned in Subsection 1.1, namely developing a theory of integration which encompasses the Riemann theory. Further, the collection of integrable functions is a complete metric space under the metric induced by Lebesgue integration, thereby accomplishing the second goal as well. This is simply the completeness of $L^1\left([a,b], \mathcal{L}([a,b]), \lambda\right)$, which is a consequence of Theorem 3.5.6.

4 Absolute continuity

4.1 The Radon-Nikodym theorem

Definition 4.1.1. *If ν and μ are measures on $(\Omega, \mathcal{A})$ such that $\mu(A) = 0$ implies that $\nu(A) = 0$ for all $A \in \mathcal{A}$, then ν is* absolutely continuous with respect to μ*, written as $\nu \ll \mu$.*

Definition 4.1.2. *Two measures μ and ν on $(\Omega, \mathcal{A})$ are* singular with respect to each other *or* mutually singular *if there exists $S \in \mathcal{A}$ such that*

$$\mu(S) = 0 = \nu(S^c)\,.$$

Suppose that $(\Omega, \mathcal{A})$ is a measurable space. All sets talked about are those in $\mathcal{A}$ and functions are $\mathcal{A}$-measurable functions, unless mentioned otherwise.

Definition 4.1.3. *A function $\psi : \mathcal{A} \to \mathbb{R}$ is an* additive set function *or a* signed measure *if the following is true. Whenever $A_1, A_2, \ldots$ is a finite or infinite sequence of disjoint sets in $\mathcal{A}$, it holds that*

$$\sum_n |\psi(A_n)| < \infty \tag{4.1.1}$$

and

$$\psi\left(\bigcup_n A_n\right) = \sum_n \psi(A_n)\,. \tag{4.1.2}$$

Remark 4.1.1. *1. Although* (4.1.1) *has been mentioned separately to ensure that the infinite sum is unchanged by rearrangements, a moment's thought would reveal that it is implied by* (4.1.2)*.*
2. It is worth emphasizing that the co-domain of an additive set function or a signed measure is $\mathbb{R}$ *and not* $\overline{\mathbb{R}}$*, that is,* $\pm\infty$ *have been left out. Otherwise the sum on the right hand side of* (4.1.2) *could have been undefined even for a finite collection of sets.*

Theorem 4.1.1 (Hahn decomposition)**.** *For any additive set function* ψ*, there exist disjoint sets* A^+ *and* A^- *such that* $A^+ \cup A^- = \Omega$*,* $\psi(E) \geq 0$ *for all* $E \subset A^+$ *and* $\psi(E) \leq 0$ *for all* $E \subset A^-$*.*

The proof will use the following exercise, which follows from arguments very similar to those in the proof of Theorem 2.2.1.

Exercise 4.1.1. *Show that if* $E_n \uparrow E$ *or* $E_n \downarrow E$*, then* $\psi(E_n) \to \psi(E)$*.*

Proof of Theorem 4.1.1. Let $\alpha = \sup\{\psi(A) : A \in \mathcal{A}\}$. Our first claim is to show that the supremum is achieved, that is, there exists $A^+ \subset \Omega$ with

$$\psi(A^+) = \alpha\,. \tag{4.1.3}$$

Let $A_n \in \mathcal{A}$ be such that

$$\lim_{n\to\infty} \psi(A_n) = \alpha\,.$$

For $n \geq 1$, let $B_n \in \sigma(A_1, \ldots, A_n)$ be such that

$$\psi(B_n) = \max\{\psi(E) : E \in \sigma(A_1, \ldots, A_n)\}\,.$$

Clearly, $\psi(B_n) \geq \psi(A_n)$.

We claim that if $E \subset B_n$ and $E \in \sigma(A_1, \ldots, A_n)$, then $\psi(E) \geq 0$. If not, then

$$\psi(B_n \setminus E) = \psi(B_n) - \psi(E) > \psi(B_n)\,,$$

contradicting the choice of B_n. Thus, $\psi(E) \geq 0$ for such an E. For positive integers $m < n$,

$$(B_m \cup \ldots \cup B_n) \setminus (B_m \cup \ldots \cup B_{n-1}) \subset B_n$$

and the LHS is an element of $\sigma(A_1, \ldots, A_n)$. Therefore, ψ evaluated on the LHS is non-negative, which shows that

$$\psi(B_m \cup \ldots \cup B_{n-1}) \leq \psi(B_m \cup \ldots \cup B_n)\,.$$

Proceeding inductively, it can be shown that

$$\psi(B_m) \leq \psi(B_m \cup \ldots \cup B_n),\ m < n\,.$$

In view of Exc 4.1.1,

$$\psi\left(\bigcup_{n=m}^{\infty} B_n\right) = \lim_{n\to\infty} \psi(B_m \cup \ldots \cup B_n) \geq \psi(B_m) \geq \psi(A_m)\,.$$

Setting

$$A^+ := \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} B_n\,,$$

it follows that

$$\psi(A^+) = \lim_{m\to\infty} \psi\left(\bigcup_{n=m}^{\infty} B_n\right) \geq \lim_{m\to\infty} \psi(A_m) = \alpha\,,$$

proving (4.1.3).

The above shows that, in particular, $\alpha < \infty$. If $E \subset A^+$ and $\psi(E) < 0$, then

$$\psi(A^+ \setminus E) > \psi(A^+)$$

which is impossible. If $E \subset \Omega \setminus A^+ =: A^-$, then by similar arguments it follows that $\psi(E) \leq 0$. This completes the proof. □

Lemma 4.1.1. *Suppose that the finite measures μ and ν on $(\Omega, \mathcal{A})$ are not mutually singular. Then, there exists A with $\mu(A) > 0$ and $\varepsilon > 0$ such that $\varepsilon\mu(E) \leq \nu(E)$ for all $E \subset A$.*

Proof. For $n \geq 1$, let $A_n^+ \cup A_n^-$ be the Hahn decomposition of the additive set function $\nu - n^{-1}\mu$. Set

$$M = \bigcup_n A_n^+\,.$$

Thus, $M^c \subset A_n^-$ for every n, and hence

$$\nu(M^c) \leq n^{-1}\mu(M^c)\,,$$

showing that $\nu(M^c) = 0$. The fact that ν and μ are not mutually singular implies that $\mu(M) > 0$. Therefore, $\mu(A_n^+) > 0$ for some n. Set $A := A_n^+$ and $\varepsilon := n^{-1}$. □

Theorem 4.1.2. *Let μ, ν be σ-finite measures on $(\Omega, \mathcal{A})$. There exists a measurable function $f : \Omega \to [0, \infty)$ and a measure ν_s singular with respect to μ such that*

$$\nu(A) = \int_A f d\mu + \nu_s(A), \text{ for all } A \in \mathcal{A}\,.$$

Proof. Suppose first that ν and μ are finite measures. Let $\mathcal{G}$ be the class of measurable functions $g : \Omega \to [0, \infty)$ such that

$$\int_E g d\mu \le \nu(E) \text{ for all } E\,.$$

If $g, g' \in \mathcal{G}$, then $g \vee g' \in \mathcal{G}$ because

$$\begin{aligned}\int_E g \vee g' d\mu &= \int_{E\cap[g\ge g']} g d\mu + \int_{E\cap[g<g']} g' d\mu \\ &\le \nu(E)\,.\end{aligned}$$

The monotone convergence theorem shows that $\mathcal{G}$ is closed under increasing limits, and therefore under countable supremum.

Set

$$\alpha := \sup_{g\in\mathcal{G}} \int g d\mu\,.$$

Let $g_n \in \mathcal{G}$ be such that

$$\lim_{n\to\infty} \int g_n d\mu = \alpha\,.$$

Set

$$f = \sup_n g_n\,.$$

Thus, $f \in \mathcal{G}$ and $\int f d\mu = \alpha$.

Set

$$\bar{\nu}(E) := \nu(E) - \int_E f d\mu,\ E \in \mathcal{A}\,.$$

Clearly, $\bar{\nu}$ is a measure. All that needs to be shown is $\bar{\nu}$ is singular with respect to μ.

Suppose that is not the case. By Lemma 4.1.1, it follows that there exists $\varepsilon > 0$ and A with $\mu(A) > 0$ such that $\varepsilon\mu(E) \le \bar{\nu}(E)$ for all $E \subset A$. Then, for every $E \in \mathcal{A}$,

$$\begin{aligned}\int_E (f + \varepsilon \mathbf{1}_A) d\mu &= \int_E f d\mu + \varepsilon\mu(E \cap A) \\ &\le \int_E f d\mu + \bar{\nu}(E \cap A) \\ &\le \int_E f d\mu + \bar{\nu}(E) \\ &= \nu(E)\,.\end{aligned}$$

Therefore, $f + \varepsilon\mathbf{1}_A \in \mathcal{G}$. Notice that

$$\int (f + \varepsilon\mathbf{1}_A) d\mu = \alpha + \varepsilon\mu(A) > \alpha\,,$$

which is a contradiction. This shows that $\bar{\nu}$ and μ are mutually singular, and thus completes the proof for finite measures.

Now suppose that μ and ν are σ-finite measures on $(\Omega, \mathcal{A})$. Then there exist disjoint sets $A_1, A_2, \ldots$ such that

$$\Omega = \bigcup_{n=1}^{\infty} A_n \,,$$

and

$$\mu(A_n) + \nu(A_n) < \infty,\, n \geq 1\,.$$

For $n \geq 1$ and $A \in \mathcal{A}$, define

$$\begin{aligned} \mu_n(A) &:= \mu(A \cap A_n)\,, \\ \nu_n(A) &:= \nu(A \cap A_n)\,. \end{aligned}$$

Then, μ_n and ν_n are finite measures for which the result is true. Therefore, there exists a measurable $f_n : \Omega \to [0, \infty)$ and a measure ν_s^n which is singular with respect to μ_n such that

$$\nu_n(A) = \int_A f_n d\mu_n + \nu_s^n(A) \text{ for all } A \in \mathcal{A}\,.$$

Therefore, for all $A \in \mathcal{A}$,

$$\begin{aligned} \nu(A) - \sum_n \nu_s^n(A) &= \sum_n \int_A f_n d\mu_n \\ &= \sum_n \int_A f_n \mathbf{1}_{A_n} d\mu = \int_A \sum_n f_n \mathbf{1}_{A_n} d\mu\,. \end{aligned}$$

Define

$$f := \sum_n f_n \mathbf{1}_{A_n}\,,$$

and the measure

$$\nu_s(A) := \sum_n \nu_s^n(A),\, A \in \mathcal{A}\,.$$

It is immediate that $0 \leq f < \infty$. The proof would follow if it can be shown that ν_s is singular with respect to μ. To that end, for $n \geq 1$, there exists $S_n \subset A_n$ such that

$$\nu_s^n(S_n) = 0 = \mu_n(S_n^c)\,.$$

Define

$$S := \bigcup_n S_n\,,$$

and note that

$$\begin{aligned}\nu_s(S) &= \sum_m \sum_n \nu_s^m(S_n) \\ &= \sum_n \nu_s^n(S_n) + \sum_{m\neq n} \nu_s^m(S_n) \\ &= \sum_{m\neq n} \nu_s^m(S_n) \\ &\leq \sum_{m\neq n} \nu_m(A_n) \\ &= 0\,.\end{aligned}$$

Furthermore,

$$\begin{aligned}\mu(S^c) &= \sum_n \mu_n(S^c) \\ &\leq \sum_n \mu_n(S_n^c) \\ &= 0\,.\end{aligned}$$

This establishes the mutual singularity of ν_s and μ, thereby completing the proof. ☐

Theorem 4.1.3 (Radon-Nikodym theorem)**.** *If μ and ν are σ-finite measures on $(\Omega, \mathcal{A})$ such that $\nu \ll \mu$, then there exists a measurable function $f : \Omega \to [0, \infty)$ such that*

$$\nu(A) = \int_A f d\mu$$

for all $A \in \mathcal{A}$. If f and g both satisfy the above, then $\mu[f \neq g] = 0$.

Exercise 4.1.2. *If f, g are non-negative measurable functions on a σ-finite measure space $(\Omega, \mathcal{A}, \mu)$ such that*

$$\int_A f d\mu = \int_A g\, d\mu, \text{ for all } A \in \mathcal{A}\,,$$

then show that $f = g$ almost everywhere.

Proof of Theorem 4.1.3. By Theorem 4.1.2, there exists a measurable function $f : \Omega \to [0, \infty)$ and a measure ν_s singular with respect to μ such that

$$\nu(A) = \int_A f d\mu + \nu_s(A),\ A \in \mathcal{S}\,.$$

The mutual singularity implies the existence of S such that

$$\nu_s(S) = 0 = \mu(S^c)\,.$$

Since $\nu \ll \mu$, it follows that

$$0 = \nu(S^c) \geq \nu_s(S^c)\,.$$

Thus, ν_s is the null measure which completes the proof. $\square$

Corollary 4.1.1. *For σ-finite measures μ, ν on $(\Omega, \mathcal{A})$, there exists a measurable $f : \Omega \to [0, \infty)$ such that*

$$\nu(A) = \int_A f d\mu,\ A \in \mathcal{S}\,, \tag{4.1.4}$$

if and only if $\nu \ll \mu$.

Definition 4.1.4. *The function f is the* Radon-Nikodym derivative *or* density *of ν with respect to μ. We write*

$$f = \frac{d\nu}{d\mu}\,,$$

$$d\nu = f\,d\mu\,,$$

or

$$\nu(dx) = f(x)\mu(dx),\ x \in \Omega\,.$$

Theorem 4.1.4 (Lebesgue decomposition)**.** *Let μ be a σ-finite measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Then there is a decomposition*

$$\nu = \nu_{ac} + \nu_s\,,$$

such that ν_{ac} and ν_s are respectively absolutely continuous and singular with respect to the Lebesgue measure.

Proof. Follows from Theorem 4.1.1. $\square$

Definition 4.1.5. *For $-\infty \leq a < b \leq \infty$, a function $F : [a, b] \cap \mathbb{R} \to \mathbb{R}$ is* absolutely continuous *if the following holds. Given $\varepsilon > 0$, there exists $\delta > 0$ such that whenever $a_1, b_1, \ldots, a_n, b_n \in \mathbb{R}$ are such that $a \leq a_1 \leq b_1 \leq \ldots \leq a_n \leq b_n \leq b$ and*

$$\sum_{i=1}^n (b_i - a_i) \leq \delta,$$

for some $n \geq 1$, it holds that

$$\sum_{i=1}^n |F(b_i) - F(a_i)| \leq \varepsilon\,.$$

A trivial observation is that an absolutely continuous function is uniformly continuous.

Theorem 4.1.5. *Suppose $F : \mathbb{R} \to \mathbb{R}$ is a non-decreasing and continuous function such that*

$$\lim_{x\to-\infty} F(x) \text{ and } \lim_{x\to\infty} F(x)$$

are finite. Let μ be the Riemann-Stieltjes measure of F. Then F is an absolutely continuous function if and only if $\mu \ll \lambda$ where λ is the Lebesgue measure.

The proof uses the following couple of lemmas, the proofs of both of which are left as exercises.

Lemma 4.1.2. *A Borel subset A of $\mathbb{R}$ has Lebesgue measure zero if and only if the following holds. Given $\varepsilon > 0$, there exist $a_1, b_1, a_2, b_2, \ldots \in \mathbb{R}$ such that $a_n < b_n$ for all n,*

$$(a_m, b_m) \cap (a_n, b_n) = \emptyset \text{ if } m \neq n\,,$$

$$A \subset \bigcup_{n=1}^{\infty} (a_n, b_n)\,,$$

and

$$\sum_{n=1}^{\infty} (b_n - a_n) \leq \varepsilon\,.$$

Proof. Exercise. □

Lemma 4.1.3. *Consider any measure space $(\Omega, \mathcal{A}, \mu)$ and an integrable function f on Ω. Given $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$\int_A |f|\, d\mu \leq \varepsilon \text{ whenever } A \in \mathcal{A},\ \mu(A) \leq \delta\,.$$

Proof. Exercise. □

Proof of Theorem 4.1.5. We start with proving the "only if" part, that is, we assume F is an absolutely continuous function. Fix a Borel set A with $\lambda(A) = 0$. We need to show that $\mu(A) = 0$. Fix an arbitrary $\varepsilon > 0$. The assumed hypothesis that F is absolutely continuous implies there exists $\delta > 0$ such that $a_1 \leq b_1 \leq \ldots \leq a_n \leq b_n$ and

$$\sum_{i=1}^{n} (b_i - a_i) \leq \delta\,,$$

implies that

$$\sum_{i=1}^{n} [F(b_i) - F(a_i)] \leq \varepsilon,$$

since F is non-decreasing. A moment's thought reveals that by considering a permutation of $\{1, \ldots, n\}$, we can write

$$\sum_{i=1}^{n} [F(b_i) - F(a_i)] \le \varepsilon \tag{4.1.5}$$

whenever $a_i < b_i$, the intervals (a_i, b_i) are disjoint and $\sum_{i=1}^{n}(b_i - a_i) \le \delta$. That is, $a_1 \le \ldots \le a_n$ is not needed for (4.1.5) to hold.

As $\lambda(A) = 0$, Lemma 4.1.2 implies the existence of $a_1, b_1, \ldots$ such that $a_n < b_n$,

$$(a_m, b_m) \cap (a_n, b_n) = \emptyset \text{ if } m \ne n\,,$$

$$A \subset \bigcup_{n=1}^{\infty} (a_n, b_n)\,,$$

and

$$\sum_{n=1}^{\infty} (b_n - a_n) \le \delta\,.$$

For a fixed n,

$$\mu\left(\bigcup_{k=1}^{n} (a_k, b_k)\right) \le \sum_{k=1}^{n} \mu\big((a_k, b_k]\big) = \sum_{k=1}^{n} (F(b_k) - F(a_k)) \le \varepsilon\,,$$

the rightmost inequality following from the choice of δ and (4.1.5). Letting $n \to \infty$ and using the continuity of measure from below, it follows that

$$\varepsilon \ge \mu\left(\bigcup_{k=1}^{\infty} (a_k, b_k)\right) \ge \mu(A)\,.$$

Since ε is arbitrary, $\mu(A) = 0$ follows. This proves the "only if" part.

Conversely, suppose that $\mu \ll \lambda$. The assumptions that F has finite limits at ±∞ imply $\mu(\mathbb{R}) < \infty$. Theorem 4.1.3, which is the Radon-Nikodym theorem, implies there exists $f : \mathbb{R} \to [0, \infty)$ Borel measurable such that

$$\mu(A) = \int_A f(x)\, dx \text{ for all } A \in \mathcal{B}(\mathbb{R})\,.$$

Since $\mu(\mathbb{R}) < \infty$, it follows that f is integrable.

Fix $\varepsilon > 0$. Lemma 4.1.3 implies the existence of $\delta > 0$ such that

$$\lambda(A) \le \delta \Rightarrow \int_A f(x)\, dx \le \varepsilon\,.$$

Let $a_1 \le b_1 \le \ldots \le a_n \le b_n$ be such that

$$\sum_{i=1}^{n} (b_i - a_i) \le \delta\,.$$

Defining $A = (a_1, b_1] \cup \ldots \cup (a_n, b_n]$, the above simply becomes $\lambda(A) \le \delta$. The choice of δ implies

$$\varepsilon \ge \int_A f(x)\,dx = \mu(A) = \sum_{i=1}^n [F(b_i) - F(a_i)] = \sum_{i=1}^n |F(b_i) - F(a_i)| \ .$$

Absolute continuity of F thus follows, which completes the proof. □

The following result is a trivial corollary of Theorem 4.1.5.

Theorem 4.1.6. *Suppose* $F : [a, b] \to \mathbb{R}$ *is non-decreasing for some* $-\infty < a < b < \infty$*. There exists an integrable* $f : [a, b] \to [0, \infty)$ *such that*

$$F(x) - F(a) = \int_a^x f(t)\,dt,\ a \le x \le b$$

if and only if F *is absolutely continuous.*

Proof. For the 'only if' part, suppose such f exists. Then F is continuous by Exc 3.6.1. Lemma 4.1.3 and arguments similar to those in the proof of Theorem 4.1.5 show F is an absolutely continuous function.

For the 'if' part, assume F is absolutely continuous on $[a, b]$. Extend F to whole of $\mathbb{R}$ by defining

$$F(x) = \begin{cases} F(a), & x < a, \\ F(b), & x > b\,. \end{cases}$$

Although F has been assumed to be absolutely continuous on $[a, b]$, a moment's thought reveals that due to the above extension, F is absolutely continuous on whole of $\mathbb{R}$.

Thus F satisfies the hypothesis of Theorem 4.1.5, that is, it is absolutely continuous on $\mathbb{R}$ and has finite limits at $\pm\infty$. Proceeding as in its proof, an $f : \mathbb{R} \to [0, \infty)$ which is the Radon-Nikodym derivative of the Riemann-Stieltjes measure μ of F with respect to Lebesgue can be obtained. Thus for $a \le x \le b$,

$$F(x) - F(a) = \mu\big((a, x]\big) = \int_a^x f(t)\,dt\,.$$

Hence the proof follows. □

4.2 Functions of Bounded Variation

This subsection is devoted to studying the Radon-Nikodym theorem and its implications for signed measures, which are as in Definition 4.1.3. While a non-decreasing function is naturally associated with a Radon measure on $\mathbb{R}$, the corresponding function for a signed measure is a "function of bounded variation" which is defined below.

Definition 4.2.1. *For* $-\infty < a \le b < \infty$ *and a function* $f : [a,b] \to \mathbb{R}$*, its* total variation *on* $[a,b]$ *is defined by*

$$V(f,[a,b]) = \sup\{v(f,P) : \, P \text{ a partition of } [a,b]\}\,,$$

where for any partition $P = (x_0, x_1, \ldots, x_n)$ *of* $[a,b]$ *with* $a = x_0 \le x_1 \le \ldots \le x_n = b$*,*

$$v(f,P) = \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \; .$$

We say f *is a* function of bounded variation *if*

$$V(f,[a,b]) < \infty\,.$$

Let $-\infty < a < b < \infty$ be fixed for this subsection. The following is a trivial exercise.

Exercise 4.2.1. *Suppose that* $P = (x_0, \ldots, x_m)$ *and* $Q = (y_1, \ldots, y_n)$ *are partitions of* $[a,b]$ *such that* Q *is a refinement of* P*, that is, for all* $i = 1, \ldots, m$*, there exists* $j \in \{1, \ldots, n\}$ *such that* $x_i = y_j$*. Show that for any* $f : [a,b] \to \mathbb{R}$*,*

$$v(f,P) \le v(f,Q)\,.$$

The following theorem characterizes functions of bounded variation.

Theorem 4.2.1. *A function* $f : [a,b] \to \mathbb{R}$ *is of bounded variation if and only if there exist non-decreasing functions* F *and* G *on* $[a,b]$ *satisfying*

$$f(x) = F(x) - G(x),\; x \in [a,b]\,.$$

Proof. For the 'if' part, suppose $f = F - G$ for non-decreasing functions F and G on $[a,b]$. For any partition $P = (x_0, \ldots, x_n)$ of $[a,b]$,

$$\begin{aligned}
v(f,P) &= \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \\
&\le \sum_{i=1}^{n} |F(x_i) - F(x_{i-1})| + \sum_{i=1}^{n} |G(x_i) - G(x_{i-1})| \\
&= \sum_{i=1}^{n} [F(x_i) - F(x_{i-1})] + \sum_{i=1}^{n} [G(x_i) - G(x_{i-1})] \\
&= F(b) - F(a) + G(b) - G(a),
\end{aligned}$$

the equality in the penultimate line holds because F, G are non-decreasing. Thus

$$V(f,[a,b]) \le F(b) - F(a) + G(b) - G(a) < \infty,$$

showing that f is a function of bounded variation.

Conversely, suppose that $V(f,[a,b])<\infty$. Define

$$F(x)=V(f,[a,x]),\; x\in[a,b]\,.$$

Our first claim is that

$$F(y)-F(x)=V(f,[x,y]),\; a\le x<y\le b\,. \tag{4.2.1}$$

Fix x,y as above. Both $F(x)$ and $V(f,[x,y])$ are defined in terms of supremum, the former being over all partitions of $[a,x]$ and the latter over those of $[x,y]$. Thus

$$\begin{aligned}
&F(x)+V(f,[x,y])\\
&=\sup\{v(f,P):\,P\text{ partition of }[a,x]\}+\sup\{v(f,Q):\,Q\text{ partition of }[x,y]\}\\
&=\sup\{v(f,P)+v(f,Q):\,P,Q\text{ respective partitions of }[a,x],[x,y]\}\\
&\le\sup\{v(R):\,R\text{ partition of }[a,y]\}\\
&=F(y)\,.
\end{aligned}$$

In other words,

$$F(y)-F(x)\ge V(f,[x,y])\,. \tag{4.2.2}$$

For the reverse inequality, let P be any partition of $[a,y]$. Consider a refinement P' of P obtained by adding the partition point x. Thus P' is essentially the joining of two partitions, say Q and R, of $[a,x]$ and $[x,y]$, respectively. Exercise 4.2.1 shows

$$\begin{aligned}
v(f,P)&\le v(f,P')\\
&=v(f,Q)+v(f,R)\\
&\le V(f,[a,x])+V(f,[x,y])\\
&=F(x)+V(f,[x,y])\,.
\end{aligned}$$

Since this holds for every partition P of $[a,y]$, we get that

$$F(y)\le F(x)+V(f,[x,y]),$$

which together with (4.2.2) establishes (4.2.1).

Finally, let $G=F-f$. A trivial consequence of (4.2.1) is that F is non-decreasing. For $a\le x<y\le b$, (4.2.1) again implies

$$F(y)-F(x)\ge|f(y)-f(x)|\ge f(y)-f(x),$$

which implies that G is also a non-decreasing function. Thus f is the difference of two non-decreasing functions and hence the ‘only if’ part follows. This completes the proof. ☐

The following exercise in real analysis will be needed for constructing a signed measure from a right continuous function of bounded variation.

Exercise 4.2.2. *Suppose $F : [a,b] \to \mathbb{R}$ is non-decreasing.*

1. *Show that for all $x \in (a,b]$, the left limit of F at x, denoted by $F(x-)$, exists and similarly for all $x \in [a,b)$, the right limit of F at x, denoted by $F(x+)$, exists.*
2. *Show that the set of discontinuities of F is countable.*
3. *Let $\{x_1, x_2, \ldots\}$ be an enumeration of $\{x \in [a,b) : F(x) \neq F(x+)\}$ which is the set of points at which F is not right continuous. Show that*
$$\sum_{n\ge1} [F(x_n+) - F(x_n)] \le F(b) - F(a) < \infty\,.$$
4. *Define*
$$\xi(x) = \sum_{n\ge1 : x_n < x} [F(x_n+) - F(x_n)]\,,\ a \le x \le b\,. \tag{4.2.3}$$
Show that ξ is a non-decreasing left continuous function.
5. *Show that $F - \xi$ is a non-decreasing right continuous function.*

Example 4.2.1. *Consider*

$$F(x) = \begin{cases} 0, & a \le x < \frac{a+b}{2}, \\ 1, & x = \frac{a+b}{2}, \\ 2, & \frac{a+b}{2} < x \le b\,. \end{cases}$$

For this example, ξ as in (4.2.3) turns out to be the following:

$$\xi(x) = \begin{cases} 0, & a \le x \le \frac{a+b}{2}, \\ 1, & \frac{a+b}{2} < x \le b\,. \end{cases}$$

Thus ξ is left continuous (but not right continuous). Further,

$$F(x) - \xi(x) = \begin{cases} 0, & a \le x < \frac{a+b}{2}, \\ 1, & \frac{a+b}{2} \le x \le b, \end{cases}$$

showing that $F - \xi$ is right continuous.

The following result is a step towards obtaining a signed measure from a right continuous function of bounded variation.

Theorem 4.2.2. *Suppose $f : [a,b] \to \mathbb{R}$ is a function of bounded variation. There exist non-decreasing F, G on $[a,b]$ such that $f = F - G$ and*

$$\{x \in [a,b) : F(x) \neq F(x+)\} \cap \{x \in [a,b) : G(x) \neq G(x+)\} = \emptyset, \tag{4.2.4}$$

that is, for all $x \in [a,b)$, either F or G is continuous at x.

Proof. Theorem 4.2.1 implies the existence of non-decreasing functions $\overline{F}$ and $\overline{G}$ such that $f = \overline{F} - \overline{G}$. The idea of the proof is to look at point where both $\overline{F}$ and $\overline{G}$ fail to be right continuous. At such points, both $\overline{F}$ and $\overline{G}$ have a right jump. The jump with the larger magnitude will be used to compensate for the smaller one. This would ensure that at least one of $\overline{F}$ and $\overline{G}$ becomes right continuous at that point. This is made precise in the following way.

Let $\{x_1, x_2, \ldots\}$ be an enumeration of the set

$$\{x \in [a,b) : \text{ Either of } \overline{F}, \overline{G} \text{ is not right continuous at } x\}$$

which is countable by Exercise 4.2.2. Define

$$\alpha(x) = \big(\overline{F}(x+) - \overline{F}(x)\big) - \big(\overline{G}(x+) - \overline{G}(x)\big),\ a \le x < b,$$

and

$$F(x) = \overline{F}(x) - \sum_{n: x_n < x} \left(\big[\overline{F}(x_n+) - \overline{F}(x_n)\big] - (\alpha(x_n))^+ \right), \tag{4.2.5}$$

$$G(x) = \overline{G}(x) - \sum_{n: x_n < x} \left(\big[\overline{G}(x_n+) - \overline{G}(x_n)\big] - (\alpha(x_n))^- \right), \tag{4.2.6}$$

for all $a \le x \le b$, where $z^+ = z \vee 0$ and $z^- = (-z) \vee 0$ for all $z \in \mathbb{R}$, as usual.

It is immediate that for all $x \in [a,b]$,

$$\begin{aligned}
&F(x) - G(x) \\
&= \overline{F}(x) - \overline{G}(x) \\
&\quad - \sum_{n: x_n < x} \big[\overline{F}(x_n+) - \overline{F}(x_n) - \overline{G}(x_n+) + \overline{G}(x_n) - (\alpha(x_n))^+ + (\alpha(x_n))^-\big] \\
&= \overline{F}(x) - \overline{G}(x) - \sum_{n: x_n < x} \big[\overline{F}(x_n+) - \overline{F}(x_n) - \overline{G}(x_n+) + \overline{G}(x_n) - \alpha(x_n)\big] \\
&= \overline{F}(x) - \overline{G}(x) \\
&= f(x),
\end{aligned}$$

the equality in the penultimate line following from the definition of $\alpha(x_n)$.

Rewrite

$$F(x) = \overline{F}(x) - \sum_{n: x_n < x} \big[\overline{F}(x_n+) - \overline{F}(x_n)\big] + \sum_{n: x_n < x} (\alpha(x_n))^+ .$$

Exercise 4.2.2 shows that

$$\overline{F}(x) - \sum_{n: x_n < x} \big[\overline{F}(x_n+) - \overline{F}(x_n)\big]$$

is a non-decreasing right continuous function of x, and

$$\sum_{n: x_n < x} (\alpha(x_n))^+$$

is non-decreasing in x. Thus F is non-decreasing. Furthermore,

$$\{x \in [a,b) : F(x) \neq F(x+)\} = \{x_n : \alpha(x_n) > 0\}\,.$$

Similarly, G is non-decreasing and

$$\{x \in [a,b) : G(x) \neq G(x+)\} = \{x_n : \alpha(x_n) < 0\}\,.$$

Thus (4.2.4) follows. Hence F and G have all the claimed properties. This completes the proof. ☐

Suppose $f : [a,b] \to \mathbb{R}$ is a right continuous function of bounded variation. Then F and G obtained from the above theorem are right continuous. Indeed, (4.2.4) implies that for $x \in [a,b)$, the moment either of F or G is right continuous at x, the right continuity of f implies that so is the other one at x as well. In other words, an immediate consequence of Theorem 4.2.2 is the following.

Corollary 4.2.1. *If $f : [a,b] \to \mathbb{R}$ is a right continuous function of bounded variation, then there exist non-decreasing right continuous functions F and G such that $f = F - G$.*

Recall Definition 4.1.3, in which a signed measure is defined. The above corollary helps in associating a signed measure to a right continuous function of bounded variation. This is done in the next result.

Theorem 4.2.3. *1. If $f : [a,b] \to \mathbb{R}$ is a right continuous function of bounded variation, then a unique signed measure μ on $\big((a,b], \mathcal{B}((a,b])\big)$ exists, such that*

$$\mu\big((\alpha,\beta]\big) = f(\beta) - f(\alpha) \text{ for all } a \leq \alpha \leq \beta \leq b\,. \tag{4.2.7}$$

2. If μ is a signed measure on $\big(\mathbb{R}, \mathcal{B}(\mathbb{R})\big)$, and $f[a,b] \to \mathbb{R}$ is defined by

$$f(x) = \mu\left((a,x]\right),\ x \in [a,b], \tag{4.2.8}$$

then (4.2.7) *holds and f is a right continuous function of bounded variation.*

Proof. 1. Corollary 4.2.1 implies that there exist right continuous non-decreasing F and G on $[a,b]$ such that $f = F - G$. Extend F to $\mathbb{R}$ by

$$F(x) = \begin{cases} F(a), & x < a, \\ F(b), & x > b. \end{cases}$$

Thus F is a right continuous non-decreasing function on $\mathbb{R}$.

Let μ_F be the Riemann-Stieltjes measure of F. Then

$$\mu_F\big((-\infty,a]\big) = F(a) - F(-\infty) = 0$$

and similarly

$$\mu_F\big((b,\infty)\big) = F(\infty) - F(b) = 0\,.$$

Thus $\mu_F\big((a,b]^c\big) = 0$. In other words, μ_F is a measure on $(a,b]$ satisfying

$$\mu_F\big((\alpha,\beta]\big) = F(\beta) - F(\alpha),\ a \le \alpha \le \beta \le b\,.$$

Likewise, a measure μ_G exists such that the above holds with 'F' replaced by 'G'. Letting

$$\mu(A) = \mu_F(A) - \mu_G(A),\ A \in \mathcal{B}\big((a,b]\big),$$

and recalling $f = F - G$, it is immediate that μ is a signed measure on $(a,b]$ satisfying (4.2.7).

For uniqueness, suppose μ' is another measure on $(a,b]$ such that (4.2.7) holds with μ replaced by μ'. Let

$$\mathcal{G} = \big\{A \in \mathcal{B}\big((a,b]\big) : \mu(A) = \mu'(A)\big\}\ .$$

It is immediate that

$$\mathcal{S} = \{(\alpha,\beta] : a \le \alpha \le \beta \le b\} \subset \mathcal{G}\,.$$

Since $\mathcal{S}$ is a semi-field and μ is additive, $\mathcal{G}$ contains the field generated by $\mathcal{S}$. Exercise 4.1.1 shows that $\mathcal{G}$ is a monotone class. It follows from the monotone class theorem that $\mathcal{G} = \mathcal{B}\big((a,b]\big)$ which shows $\mu = \mu'$. in other words, uniqueness follows.

2. Suppose μ is a signed measure on $\big(\mathbb{R},\mathcal{B}(\mathbb{R})\big)$ and f is defined by (4.2.8). Additivity of μ immediately shows (4.2.7). If $x_n \in [a,b]$ and $x_n \downarrow x$, then

$$(x,x_n] \downarrow \emptyset\,.$$

Hence

$$0 = \mu(\emptyset) = \lim_{n\to\infty} \mu\big((x,x_n]\big) = -f(x) + \lim_{n\to\infty} f(x_n)\,.$$

This shows right continuity of f at x and hence completes the proof. □

Definition 4.2.2. *For a right continuous f of bounded variation on $[a,b]$, the unique signed measure μ on $(a,b]$ satisfying* (4.2.7) *is the* signed measure induced by *f.*

Theorem 4.2.4. *An absolutely continuous function $f : [a,b] \to \mathbb{R}$ is a function of bounded variation. Further, its total variation function*

$$F(x) = V\left(f,[a,x]\right),\ x \in [a,b], \tag{4.2.9}$$

is also absolutely continuous.

Proof. Since f is absolutely continuous, there exists $k \geq 1$ such that

$$\sum_{i=1}^{n} |f(y_i) - f(x_{i=1})| \leq 1 \tag{4.2.10}$$

whenever $a \leq x_1 \leq y_1 \leq x_2 \leq y_2 \leq \ldots \leq x_n \leq y_n \leq b$ satisfy

$$\sum_{i=1}^{n} (y_i - x_i) \leq \frac{b-a}{k} . \tag{4.2.11}$$

Let P be any partition of $[a, b]$ and Q be a refinement of P such that

$$a + \frac{i}{k}(b-a) \in Q,\ i = 1, \ldots, k .$$

That is $Q = (x_0, \ldots, x_n)$ and there exist $0 = n_0 < n_1 < \ldots < n_k = n$ such that

$$x_{n_i} = a + \frac{i}{k}(b-a),\ i = 0, \ldots, k .$$

Exercise 4.2.1 shows

$$\begin{aligned} v(f, P) &\leq v(f, Q) \\ &= \sum_{i=0}^{n} |f(x_i) - f(x_{i-1})| \\ &= \sum_{j=1}^{k} \sum_{i=n_{j-1}+1}^{n_j} |f(x_i) - f(x_{i-1})| . \end{aligned}$$

The choice of k in (4.2.11) which ensures (4.2.10) and the observation

$$\sum_{i=n_{j-1}+1}^{n_j} (x_i - x_{i-1}) = \frac{b-a}{k}$$

imply

$$\sum_{i=n_{j-1}+1}^{n_j} |f(x_i) - f(x_{i-1})| \leq 1$$

for $j = 1, \ldots, k$. Thus $v(f, P) \leq k$ for every partition P of $[a, b]$, which shows f has bounded variation.

Let F be as in (4.2.9). For showing absolute continuity of F, fix $\varepsilon > 0$. Absolute continuity of f imply the existence of $\delta > 0$ such that

$$\sum_{i=1}^{n} |f(\beta_i) - f(\alpha_i)| \tag{4.2.12}$$

whenever $a \leq \alpha_1 \leq \beta_1 \leq \ldots \leq \alpha_n \leq \beta_n \leq b$ and

$$\sum_{i=1}^{n} (\beta_i - \alpha_i) \leq \delta \,.$$

Fix $a \leq x_1 \leq y_1 \leq \ldots x_n \leq y_n$ such that

$$\sum_{i=1}^{n} (y_i - x_i) \leq \delta \,.$$

Recall that (4.2.1) implies

$$\sum_{i=1}^{n} \big(F(y_i) - F(x_i)\big) = \sum_{i=1}^{n} V(f, [x_i, y_i]) \,.$$

Let $P_1, \ldots, P_n$ be partitions of $[x_1, y_1], \ldots, [x_n, y_n]$, respectively. Suppose $P_i = (\alpha_{i0}, \ldots, \alpha_{ik_i})$. Then

$$\sum_{i=1}^{n} \sum_{j=1}^{k_i} (\alpha_{ij} - \alpha_{i\,j-1}) = \sum_{i=1}^{n} (y_i - x_i) \leq \delta \,.$$

The choice of δ, which ensures (4.2.12), implies that

$$\varepsilon \geq \sum_{i=1}^{n} \sum_{j=1}^{k_i} |f(\alpha_{ij}) - f(\alpha_{i\,j-1})| = \sum_{i=1}^{n} v(f, P_i) \,.$$

As this is true for all choices of partitions $P_1, \ldots, P_n$ of $[x_1, y_1], \ldots, [x_n, y_n]$, respectively, it follows that

$$\varepsilon \geq \sum_{i=1}^{n} V(f, [x_i, y_i]) = \sum_{i=1}^{n} \big(F(y_i) - F(x_i)\big) \,.$$

Thus absolute continuity of F follows, which completes the proof. ☐

Definition 4.2.3. *Given a measurable space* $(\Omega, \mathcal{A})$ *on which* μ *is a signed measure and* ν *is a measure,* μ *is* absolutely continuous *with respect to* ν *or* $\mu \ll \nu$ *if*

$$\nu(A) = 0 \textit{ implies } \mu(A) = 0 \textit{ for all } \in \mathcal{A} \,.$$

Theorem 4.2.5 (Radon-Nikodym theorem for signed measures)**.** *Suppose* $(\Omega, \mathcal{A}, \nu)$ *is a* σ*-finite measure space and* μ *is a signed measure on* $(\Omega, \mathcal{A})$*. Then* $\mu \ll \nu$ *if and only if there exists* $f \in L^1(\Omega, \nu)$ *such that*

$$\mu(A) = \int_A f \, d\nu, \; A \in \mathcal{A} \,. \tag{4.2.13}$$

Proof. The 'if' part is trivial because the right hand side of (4.2.13) vanishes if $\nu(A) = 0$. For the converse part, assume $\mu \ll \nu$. Theorem 4.1.1 shows the existence of A_+ and A_- such that $A_+ = (A_-)^c$ and

$$\mu(A) \geq 0 \text{ for all } A \subset A_+ \text{ and } \mu(A) \leq 0 \text{ for all } A \subset A_- .$$

Define

$$\mu^+(A) = \mu(A \cap A_+) \text{ and } \mu^-(A) = -\mu(A \cap A_-),\ A \in \mathcal{A} .$$

It is immediate that $\mu^\pm$ are finite measures on $(\Omega, \mathcal{A})$ and $\mu = \mu^+ - \mu^-$. Further, $\nu(A) = 0$ implies that $\nu(A \cap A_+) = 0$ and hence $\mu(A \cap A_+) = 0$. Thus $\mu^+ \ll \nu$ and similarly $\mu^- \ll \nu$.

Let

$$g = \frac{d\mu^+}{d\nu} \text{ and } h = \frac{d\mu^-}{d\nu},$$

which exist due to Theorem 4.1.3. Since $\mu^\pm$ are finite measures, g and h are ν-integrable. Setting $f = g - h$ ensures (4.2.13). Hence the proof follows. ☐

Exercise 4.2.3. *Show that an f, which satisfies* (4.2.13)*, is unique upto sets of ν-measure zero, if exists.*

Theorem 4.2.6. *Suppose $f : [a, b] \to \mathbb{R}$ is a right continuous function of bounded variation and μ is the signed measure on $(a, b]$ induced by f. The following are equivalent:*

1. *f is an absolutely continuous function,*
2. *$\mu \ll \lambda$ where λ is the Lebesgue measure,*
3. *there exists a Lebesgue integrable function g on $[a, b]$ such that*

$$\int_a^x g(t)\, dt = f(x) - f(a),\ a \leq x \leq b .$$

Proof. We start with showing that 1. implies 2. Assume f is an absolutely continuous function. Theorem 4.2.4 shows that F, as in (4.2.9), is absolutely continuous. In the proof of Theorem 4.2.1, it has been shown that F and $G = F - f$ are non-decreasing. Further, G is absolutely continuous as well because so are f and F.

Let μ_F be the measure on $(a, b]$ satisfying

$$\mu_F\left((x, y]\right) = F(y) - F(x),\ a \leq x \leq y \leq b,$$

and μ_G is likewise defined. Theorem 4.1.5 shows that $\mu_F \ll \lambda$ and $\mu_G \ll \lambda$. Since $\mu = \mu_F - \mu_G$, it follows that $\mu \ll \lambda$. Thus 1. implies 2.

For showing 2. implies 3., assume $\mu \ll \lambda$. Theorem 4.2.5 shows that there exists a Lebesgue integrable g on $(a, b]$ such that

$$\mu(A) = \int_A g \, d\lambda \text{ for all } A \in \mathcal{B}\big((a, b]\big) .$$

Taking $A = (a, x]$ for $x \in [a, b]$ establishes 3.

Finally, assume 3. for showing that it implies 1. Let $\varepsilon > 0$. Since g is Lebesgue integrable on $(a, b]$, Lemma 4.1.3 shows that $\delta > 0$ exists such that

$$\int_A |g(x)| \, dx \leq \varepsilon \text{ whenever } \lambda(A) \leq \delta . \tag{4.2.14}$$

Fix $a \leq x_1 \leq y_1 \leq \ldots \leq x_n \leq y_n \leq b$ with $\sum_{i=1}^{n}(y_i - x_i) \leq \delta$ and take $A = (x_1, y_1] \cup \ldots (x_n, y_n]$. Thus $\lambda(A) \leq \delta$, which shows

$$\begin{aligned} \varepsilon &\geq \int_A |g(x)| \, dx \\ &= \sum_{i=1}^{n} \int_{x_i}^{y_i} |g(x)| \, dx \\ &\geq \sum_{i=1}^{n} \left| \int_{x_i}^{y_i} g(x) \, dx \right| \\ &= \sum_{i=1}^{n} |f(y_i) - f(x_i)| . \end{aligned} \tag{4.2.15}$$

Thus f is absolutely continuous, that is, 3. implies 1. This completes the proof. □

The following result generalizes Theorem 4.1.6. It should be noted that the latter is built on Theorem 4.1.5, which has been used in the proof of Theorem 4.2.6 above. The 'if' part of the corollary below follows from (4.2.14) - (4.2.15) and its 'only if' part is a consequence of Theorem 4.2.4 and the equivalence of 1. and 3. in Theorem 4.2.6.

Corollary 4.2.2. *A function $f : [a, b] \to \mathbb{R}$ is absolutely continuous if and only if there exists a Lebesgue integrable g on $[a, b]$ such that*

$$\int_a^x g(t) \, dt = f(x) - f(a) \textit{ for all } x \in [a, b] .$$

4.3 The fundamental theorems of calculus

A result which connects integration with differentiation is historically known as a fundamental theorem of calculus (FTC). The statements of the FTCs are of the following two types.

1. The derivative of the integral gives back the original function. The second part of Exercise 3.6.1, for example, is a claim of this kind.
2. The integral of the derivative of a function on an interval equals the increment of the function on that interval. In other words, a result guaranteeing (1.1.2) under some conditions falls in this category.

We shall proceed towards an FTC of the first kind above. Since such a claim is easy to show when the integrand is continuous, as done in 2. of Exercise 3.6.1, a natural way forward is to show that any integrable function can be approximated by continuous functions.

Definition 4.3.1. *A function $f : \mathbb{R} \to \mathbb{R}$ is* compactly supported *if there exists a compact set $K \subset \mathbb{R}$ such that*

$$f(x) = 0 \text{ for all } x \notin K\,.$$

Define

$$\begin{aligned} C_c(\mathbb{R}) &= \left\{ f : \mathbb{R} \to \mathbb{R} \,\middle|\, f \text{ is a compactly supported continuous function} \right\}, \\ C(\mathbb{R}) &= \left\{ f : \mathbb{R} \to \mathbb{R} \,\middle|\, f \text{ is a continuous function} \right\}. \end{aligned}$$

Exercise 4.3.1. *Show that $f : \mathbb{R} \to \mathbb{R}$ is compactly supported if and only if $f^{-1}(\mathbb{R} \setminus \{0\})$ is a bounded set.*

As usual, λ denotes the Lebesgue measure, and unless stated otherwise, $L^p(\mathbb{R})$ will mean $L^p(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$.

Theorem 4.3.1. *For $1 \le p < \infty$, $C_c(\mathbb{R})$ is dense in $L^p(\mathbb{R})$.*

Before embarking on the proof, let us note that the above fails if $p = \infty$.

Exercise 4.3.2. *Show that $C_c(\mathbb{R})$ is not dense in $L^\infty(\mathbb{R})$.*

The proof of Theorem 4.3.1 uses the following two lemmas.

Lemma 4.3.1. *If μ is a Radon measure on $\mathbb{R}$, then for all $A \in \mathcal{B}(\mathbb{R})$,*

$$\sup\left\{\mu(K) : K \subset A,\, K \text{ compact}\right\} = \mu(A) = \inf\left\{\mu(U) : U \supset A,\, U \text{ open}\right\}\,.$$

Further, if $\mu(A) < \infty$, then for all $\varepsilon > 0$, there exist compact K and open U such that $K \subset A \subset U$ and $\mu(U \setminus K) \le \varepsilon$.

Proof. Exercise. □

Lemma 4.3.2. *If $K \subset U \subset \mathbb{R}$ where K is a compact set and U is an open set, then there exists $f \in C_c(\mathbb{R})$ such that*

$$\mathbf{1}_K \le f \le \mathbf{1}_U\,.$$

Proof. Let

$$\delta = \inf \left\{ |x - y| : x \in K, y \in U^c \right\} .$$

Our first claim is that $\delta > 0$. Assume $\delta = 0$ for the sake of contradiction. That is, there exist $x_n \in K$ and $y_n \in U^c$ such that $|x_n - y_n| \to 0$. Since K is compact, $\{x_n\}$ has a convergent subsequence $\{x_{n_k}\}$. Denoting

$$x = \lim_{k \to \infty} x_{n_k},$$

compactness of K implies $x \in K$. Further, $y_{n_k} \to x$ as $k \to \infty$. Since U is open, U^c is closed and hence $x \in U^c$. Thus $x \in K \cap U^c$, which contradicts $K \subset U$. Therefore $\delta > 0$.

Define

$$f(x) = 1 - \frac{1}{\delta} \left(\delta \wedge d(x, K) \right), \, x \in \mathbb{R},$$

where

$$d(x, K) = \inf\{|x - y| : y \in K\}, \, x \in \mathbb{R} .$$

Since $d(\cdot, K)$ is a continuous function, so is f. For $x \in K$, $d(x, K) = 0$ which shows $f(x) = 1$. Hence $\mathbf{1}_K \leq f$. If $x \in U^c$, then $d(x, K) \geq \delta$ and hence $f(x) = 0$, which shows $f \leq \mathbf{1}_U$. Finally,

$$\{x : f(x) \neq 0\} = \{x : d(x, K) < \delta\} .$$

Since K is a bounded set, so is the set on the right hand side above. Exercise 4.3.1 shows f is compactly supported and hence $f \in C_c(\mathbb{R})$. This completes the proof. ☐

Proof of Theorem 4.3.1. Fix $p \in [1, \infty)$, $f \in L^p(\mathbb{R})$ and $\varepsilon > 0$. It has to be shown that $g \in C_c(\mathbb{R})$ exists such that

$$\|f - g\|_p \leq \varepsilon . \tag{4.3.1}$$

Without loss of generality, f can be assumed to be non-negative because if $g_1, g_2 \in C_c(\mathbb{R})$ are such that $\|f^+ - g_1\|_p \leq \varepsilon/2$ and $\|f^- - g_2\|_p \leq \varepsilon/2$, (4.3.1) would be ensured by $g = g_1 - g_2$. Thus (4.3.1) has to be proved for a non-negative $f \in L^p(\mathbb{R})$. Fix such an f.

Since $f \geq 0$, there exist simple functions s_n such that $0 \leq s_n \uparrow f$. As $0 \leq (f - s_n)^p \leq f^p$ and $\int_{\mathbb{R}} f(x)^p \, dx < \infty$, DCT implies that

$$\|s_n - f\|_p \to 0, \, n \to \infty .$$

Let $t_1 = s_1$ and for $n = 2, 3, \ldots$, $t_n = s_n - s_{n-1}$. Thus

$$\sum_{i=1}^{n} t_i = s_n \to f \text{ in } L^p, \, n \to \infty .$$

Further, $t_i \geq 0$ for all i. That is, t_i is a non-negative simple function. Hence for all $i = 1, 2, \ldots$,

$$t_i = \sum_{j=1}^{n_i} \alpha_{ij} \mathbf{1}_{A_{ij}},$$

for some $A_{i1}, \ldots, A_{in_i} \in \mathcal{B}(\mathbb{R})$ and $\alpha_{i1}, \ldots, \alpha_{in_i} > 0$. Therefore

$$f = \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \alpha_{ij} \mathbf{1}_{A_{ij}},$$

where the infinite sum on the right hand side converges in L^p. By a re-indexing, write

$$f = \sum_{n=1}^{\infty} \alpha_n \mathbf{1}_{A_n}, \tag{4.3.2}$$

where once again the sum converges in L^p, $\alpha_1, \alpha_2, \ldots > 0$ and $A_1, A_2, \ldots \in \mathcal{B}(\mathbb{R})$. The sets A_n are not necessarily disjoint, and hence we should not get confused into thinking f takes countably many values. Further, $\lambda(A_n) > 0$ for all n may be assumed because otherwise they may simply be dropped keeping the equality in (4.3.2), which holds in L^p, intact.

Fix N such that

$$\left\| f - \sum_{n=1}^{N} \alpha_n \mathbf{1}_{A_n} \right\|_p \leq \frac{\varepsilon}{2} \, .$$

Since $\lambda(A_n) > 0$, it is immediate that $\alpha_n < \infty$ for all n, because otherwise (4.3.2) would imply $\int_{\mathbb{R}} f^p \, d\lambda = \infty$. Similarly, $\alpha_n > 0$ implies $\lambda(A_n) < \infty$ for all n. Lemma 4.3.1 implies the existence of a compact set K_n and an open set U_n such that

$$K_n \subset A_n \subset U_n \text{ and } \lambda(U_n \setminus K_n) \leq \left(\frac{\varepsilon}{\alpha_n 2^{n+1}} \right)^p, \; n \geq 1 \, .$$

Using Lemma 4.3.2, $g_n \in C_c(\mathbb{R})$ satisfying $\mathbf{1}_{K_n} \leq g_n \leq \mathbf{1}_{U_n}$ can be obtained. Since $\mathbf{1}_{K_n} \leq \mathbf{1}_{A_n} \leq \mathbf{1}_{U_n}$, it follows that

$$|g_n - \mathbf{1}_{A_n}| \leq \mathbf{1}_{U_n} - \mathbf{1}_{K_n} = \mathbf{1}_{U_n \setminus K_n},$$

which implies

$$\begin{aligned} \|g_n - \mathbf{1}_{A_n}\|_p &\leq \left\| \mathbf{1}_{U_n \setminus K_n} \right\|_p \\ &= \left(\lambda(U_n \setminus K_n) \right)^{1/p} \\ &\leq \frac{\varepsilon}{\alpha_n 2^{n+1}} \, . \end{aligned}$$

Setting

$$g = \sum_{n=1}^{N} \alpha_n g_n,$$

it therefore follows that

$$\left\| g - \sum_{n=1}^{N} \alpha_n \mathbf{1}_{A_n} \right\|_p \le \sum_{n=1}^{N} \alpha_n \left\| g_n - \mathbf{1}_{A_n} \right\|_p \le \sum_{n=1}^{\infty} 2^{-n-1}\varepsilon = \frac{\varepsilon}{2}\,.$$

This in conjunction with the choice of N establishes (4.3.1), from which the proof follows. ☐

Definition 4.3.2. *For* $f \in L^1(\mathbb{R})$, *its* maximal function $Mf : \mathbb{R} \to \overline{\mathbb{R}}$ *is defined by*

$$Mf(x) = \sup_{0<r<\infty} \frac{1}{2r} \int_{x-r}^{x+r} |f(t)|\, dt,\ x \in \mathbb{R}\,.$$

The first question which arises is whether Mf is measurable. It will be answered by showing that Mf is "lower semicontinuous" which is defined below.

Definition 4.3.3. *A function* $g : \mathbb{R} \to \overline{\mathbb{R}}$ *is* lower semicontinuous *if for all* $\theta \in \mathbb{R}$, $g^{-1}\big((\theta, \infty]\big)$ *is an open set. Similarly,* $g : \mathbb{R} \to \overline{\mathbb{R}}$ *is* upper semicontinuous *if* $g^{-1}\big([-\infty, \theta)\big)$ *is open for all* $\theta \in \mathbb{R}$.

Theorem 4.3.2. *For* $f \in L^1(\mathbb{R})$, Mf *is lower semicontinuous.*

Proof. Let $f \in L^1(\mathbb{R})$ and $\theta \in \mathbb{R}$. Suppose $x \in [Mf > \theta]$, that is, $Mf(x) > \theta$. Thus there exists $r > 0$ such that

$$\frac{1}{2r} \int_{x-r}^{x+r} |f(t)|\, dt > \theta\,.$$

Let $\delta > 0$ be such that

$$\theta < \frac{r}{r+\delta} \frac{1}{2r} \int_{x-r}^{x+r} |f(t)|\, dt = \frac{1}{2(r+\delta)} \int_{x-r}^{x+r} |f(t)|\, dt\,. \tag{4.3.3}$$

For $y \in [r-\delta, r+\delta]$, $[y-r-\delta, y+r+\delta] \supset [x-r, x+r]$ and hence

$$\int_{y-r-\delta}^{y+r+\delta} |f(t)|\, dt \ge \int_{x-r}^{x+r} |f(t)|\, dt\,.$$

Thus

$$\frac{1}{2(r+\delta)} \int_{y-r-\delta}^{y+r+\delta} |f(t)|\, dt \ge \frac{1}{2(r+\delta)} \int_{x-r}^{x+r} |f(t)|\, dt > \theta,$$

(4.3.3) implying the rightmost inequality which is strict. Hence $Mf(y) > \theta$, which shows $[x-\delta, x+\delta] \subset [Mf > \theta]$. In other words, every point of $[Mf > \theta]$ is an interior point, showing that the set is open. As this holds for all $\theta \in \mathbb{R}$, the proof follows. ☐

An immediate consequence of the above result is that Mf is measurable.

Theorem 4.3.3. *For $f \in L^1(\mathbb{R})$ and $\theta > 0$,*

$$\lambda\left([Mf > \theta]\right) \leq \frac{3}{\theta}\|f\|_1 .$$

The proof is built on the following lemma.

Lemma 4.3.3. *For $x_1, \ldots, x_n \in \mathbb{R}$ and $r_1, \ldots, r_n > 0$, there exists $S \subset \{1, \ldots, n\}$ such that the sets in $\{(x_i - r_i, x_i + r_i) : i \in S\}$ are disjoint and*

$$\bigcup_{i=1}^{n}(x_i - r_i, x_i + r_i) \subset \bigcup_{i \in S}(x_i - 3x_i, x_i + 3r_i) .$$

Proof. Assume without loss of generality that $r_1 \geq r_2 \geq \ldots \geq r_n$. A trivial observation is that if $y, z \in \mathbb{R}$ and $0 \leq s \leq t$ are such that

$$(y - s, y + s) \cap (z - t, z + t) \neq \emptyset,$$

then

$$(y - s, y + s) \subset (z - 3t, z + 3t) .$$

Let $i_1 = 1$ and

$$i_2 = \min\{j > 1 : (x_j - r_j, x_j + r_j) \cap (x_{i_1} - r_{i_1}, x_{i_1} + r_{i_1}) = \emptyset\} .$$

Proceed inductively by letting

$$i_{k+1} = \min\{j > i_k : (x_j - r_j, x_j + r_j) \cap (x_{i_u} - r_{i_u}, x_{i_u} + r_{i_u}) = \emptyset,\ u = 1, \ldots, k\}$$

until $i_k = n$ for some $k \geq 1$ or the set on the right hand side above is empty.

Letting $S = \{i_1, \ldots, i_k\}$, it is obvious from the construction that for all $j \in \{1, \ldots, n\} \setminus S$,

$$(x_j - r_j, x_j + r_j) \cap (x_i - r_i, x_i + r_i) \neq \emptyset$$

for some $i \in S$ with $i < j$ and hence the observation above, along with $r_i \geq r_j$, implies

$$(x_j - r_j, x_j + r_j) \subset \bigcup_{i \in S : i < j}(x_i - 3r_i, x_i + 3r_i) .$$

It thus follows that

$$\bigcup_{i=1}^{n}(x_i - r_i, x_i + r_i) \subset \bigcup_{i \in S}(x_i - 3x_i, x_i + 3r_i) .$$

It is also immediate that $\{(x_i - r_i, x_i + r_i) : i \in S\}$ is a collection of disjoint sets. Hence the proof follows. ☐

Proof of Theorem 4.3.3. Let $f \in L^1(\mathbb{R})$ and $\theta > 0$. By an appeal to Lemma 4.3.1, it suffices to show that for all compact $K \subset [Mf > \theta]$,

$$\lambda(K) \le \frac{3}{\theta}\|f\|_1 \,. \tag{4.3.4}$$

Fix such a K.

For all $x \in K$, $Mf(x) > \theta$. Hence there exists $r(x) > 0$ such that

$$\frac{1}{2r(x)} \int_{x-r(x)}^{x+r(x)} |f(t)|\, dt \ge \theta \,. \tag{4.3.5}$$

Since $\{(x - r(x), x + r(x)) : x \in K\}$ is an open cover of K which is compact, there exist $x_1, \ldots, x_n \in K$ with

$$K \subset \bigcup_{i=1}^{n} (x_i - r_i, x_i + r_i),$$

where $r_i = r(x_i)$ for $i = 1, \ldots, n$.

Let $S \subset \{1, \ldots, n\}$ be obtained from the claim of Lemma 4.3.3. Thus

$$K \subset \bigcup_{i=1}^{n} (x_i - r_i, x_i + r_i) \subset \bigcup_{i \in S} (x_i - 3r_i, x_i + 3r_i),$$

a consequence of which is

$$\begin{aligned}
\lambda(K) &\le \sum_{i \in S} \lambda\big((x_i - 3r_i, x_i + 3r_i)\big) \\
&= \sum_{i \in S} 6r_i \\
&\le 3 \sum_{i \in S} \frac{1}{\theta} \int_{x_i - r_i}^{x_i + r_i} |f(t)|\, dt \\
&= \frac{3}{\theta} \int_{\bigcup_{i \in S}(x_i - r_i, x_i + r_i)} |f(t)|\, dt \le \frac{3}{\theta}\|f\|_1,
\end{aligned}$$

(4.3.5) implying the penultimate line and the equality in the last line following from the disjointness of the sets in $\{(x_i - r_i, x_i + r_i) : i \in S\}$. Hence (4.3.4) follows, which completes the proof. ☐

Definition 4.3.4. *For $f \in L^1(\mathbb{R})$, $x \in \mathbb{R}$ is a* Lebesgue point *of f if*

$$\lim_{r \downarrow 0} \frac{1}{2r} \int_{x-r}^{x+r} |f(t) - f(x)|\, dt = 0 \,.$$

The following result illustrates the significance of Lebesgue points.

Theorem 4.3.4. *Suppose $f \in L^1(\mathbb{R})$. Define*

$$F(x) = \int_{-\infty}^{x} f(t)\, dt,\; x \in \mathbb{R}\,.$$

1. If x is a Lebesgue point of f, then F is differentiable at x and

$$F'(x) = f(x)\,.$$

2. Points at which f is continuous are Lebesgue points of f.

Proof. 1. Let f and F be as given and x be a Lebesgue point of f. Then for $h > 0$,

$$\begin{aligned}\left|\frac{F(x+h)-F(x)}{h} - f(x)\right| &= \frac{1}{h}\left|\int_x^{x+h} (f(t)-f(x))\; dt\right| \\ &\le \frac{1}{h}\int_x^{x+h} |f(t)-f(x)|\, dt \\ &\le \frac{1}{h}\int_{x-h}^{x+h} |f(t)-f(x)|\, dt\,.\end{aligned}$$

Thus

$$\limsup_{h\downarrow 0}\left|\frac{F(x+h)-F(x)}{h} - f(x)\right| \le 2\limsup_{h\downarrow 0}\frac{1}{2h}\int_{x-h}^{x+h} |f(t)-f(x)|\, dt = 0,$$

the equality being a consequence of x being a Lebesgue point. Thus the right derivative of F at x equals $f(x)$. A similar argument works for the left derivative as well, which shows

$$F'(x) = f(x)\,.$$

2. Let x be a continuity point of f. Fix $\varepsilon > 0$. Continuity at x implies there exists $\delta > 0$ such that

$$|f(t) - f(x)| \le \varepsilon,\; t \in [x-\delta, x+\delta]\,.$$

Thus for $0 < r \le \delta$,

$$\frac{1}{2r}\int_{x-r}^{x+r} |f(t)-f(x)|\, dt \le \frac{1}{2r}\int_{x-r}^{x+r} \varepsilon\, dt = \varepsilon\,.$$

This shows

$$\lim_{r\downarrow 0}\frac{1}{2r}\int_{x-r}^{x+r} |f(t)-f(x)|\, dt = 0,$$

from which the proof follows. ☐

Theorem 4.3.5. *For $f \in L^1(\mathbb{R})$,*

$$\{x \in \mathbb{R} : x \textit{ is not a Lebesgue point of } f\} \in \mathcal{L}(\mathbb{R})$$

and the set on the left hand side has Lebesgue measure zero. That is, almost every real number is a Lebesgue point of f.

Proof. Define for all $h \in L^1(\mathbb{R})$ and $0 < r < \infty$, $T_r h : \mathbb{R} \to \overline{\mathbb{R}}$ by

$$T_r h(x) = \frac{1}{2r} \int_{x-r}^{x+r} |h(t) - h(x)| \, dt, x \in \mathbb{R},$$

and $Th : \mathbb{R} \to \overline{\mathbb{R}}$ by

$$Th(x) = \limsup_{r \downarrow 0} T_r h(x), \; x \in \mathbb{R} \, .$$

Fix $n \geq 1$ and let $g \in C_c(\mathbb{R})$ be such that $\|f - g\|_1 \leq \frac{1}{n}$, which is obtained from Theorem 4.3.1. Let $h = f - g$. Thus for all $x \in \mathbb{R}$,

$$T_r f(x) = (T_r(g + h))(x) \leq T_r g(x) + T_r h(x) \, .$$

Since g is continuous, every point is a Lebesgue point by the second part of Theorem 4.3.4. Taking $\limsup_{r \downarrow 0}$ we get

$$Tf(x) \leq Th(x), \; x \in \mathbb{R} \, . \tag{4.3.6}$$

For all $x \in \mathbb{R}$,

$$\begin{aligned} T_r h(x) &\leq \frac{1}{2r} \int_{x-r}^{x+r} (|h(t)| + |h(x)|) \, dt \\ &= |h(x)| + \frac{1}{2r} \int_{x-r}^{x+r} |h(t)| \, dt \\ &\leq |h(x)| + Mh(x) \, . \end{aligned}$$

Therefore,

$$Th(x) \leq |h(x)| + Mh(x) \, .$$

This in conjunction with (4.3.6) shows that

$$[Tf > 2\xi] \subset [|h| > \xi] \cup [Mh > \xi], \; \xi > 0 \, . \tag{4.3.7}$$

So far, n was fixed. Now suppose for all $n = 1, 2, \ldots$, $g_n \in C_c(\mathbb{R})$ is chosen to satisfy $\|f - g_n\|_1 \leq \frac{1}{n}$. Letting $h_n = f - g_n$, (4.3.7) implies

$$[Tf > 2\xi] \subset [|h_n| > \xi] \cup [Mh_n > \xi], \; n = 1, 2, \ldots, \; \xi > 0 \, .$$

For a fixed $\xi > 0$, it thus follows that

$$[Tf > 2\xi] \subset \bigcap_{n=1}^{\infty} \left([|h_n| > \xi] \cup [Mh_n > \xi]\right) .$$

Theorems 3.4.4 and 4.3.3 imply

$$\lambda\left([|h_n| > \xi] \cup [Mh_n > \xi]\right) \le \frac{4}{\xi}\|h_n\|_1 \le \frac{4}{n\xi} .$$

Therefore

$$\lambda\left(\bigcap_{n=1}^{\infty} \left([|h_n| > \xi] \cup [Mh_n > \xi]\right)\right) = 0 .$$

Theorem 3.6.1 implies that $[Tf > 2\xi] \in \mathcal{L}(\mathbb{R})$ and

$$\lambda\left([Tf > 2\xi]\right) = 0 .$$

As $[Tf > 2/k] \uparrow [Tf > 0]$, it follows that

$$Tf = 0 \text{ a.e.}$$

This completes the proof. ☐

Let us adopt the following notation for the rest of this subsection:

$$F'(x) = \begin{cases} \frac{d}{dx}F(x), & \text{if the derivative exists at } x, \\ 0, & \text{otherwise} . \end{cases} \tag{4.3.8}$$

Exercise 4.3.3. *Suppose $F : \mathbb{R} \to \mathbb{R}$ is a Borel function. Show that*

$$\{x \in \mathbb{R} : F \textit{ is differentiable at } x\} \in \mathcal{B}(\mathbb{R}),$$

and F', as in (4.3.8)*, is a Borel function.*

The next result is an FTC of the first kind.

Theorem 4.3.6 (FTC, version 1)**.** *Suppose $f \in L^1(\mathbb{R})$. Define*

$$F(x) = \int_{-\infty}^{x} f(t)\, dt,\ x \in \mathbb{R} .$$

Then F is differentiable a.e. and

$$F' = f \text{ a.e.}$$

Proof. For every Lebesgue point x of f, F is differentiable at x and $F'(x) = f(x)$ by the first part of Theorem 4.3.4. Since almost every point is a Lebesgue point of f by Theorem 4.3.5, the proof follows. ☐

The next result combines the above result with Corollary 4.2.2.

Theorem 4.3.7. *For* $-\infty < a < b < \infty$*,* $F : [a,b] \to \mathbb{R}$ *is an absolutely continuous function if and only if* F *is differentiable a.e. on* $[a,b]$*,*

$$\int_a^b |F'(x)|\,dx < \infty$$

and

$$\int_a^x F'(t)\,dt = F(x) - F(a) \textit{ for all } x \in [a,b]\,.$$

Proof. The 'if' part follows directly from Corollary 4.2.2. Conversely, assume F is absolutely continuous. Once again, Corollary 4.2.2 implies there exists an integrable $f : [a,b] \to \mathbb{R}$ such that

$$\int_a^x f(t)\,dt = F(x) - F(a),\ x \in [a,b]\,. \tag{4.3.9}$$

Extend f to whole of $\mathbb{R}$ by defining it to be zero outside $[a,b]$. Since

$$F(x) = F(a) + \int_{-\infty}^x f(t)\,dt,\ x \in [a,b],$$

Theorem 4.3.6 shows that F is differentiable a.e. and

$$F' = f \text{ a.e. on } [a,b]\,.$$

Thus F' is integrable on $[a,b]$ and (4.3.9) shows that

$$\int_a^x F'(t)\,dt = F(x) - F(a) \text{ for all } x \in [a,b]\,.$$

Hence the 'only if' part follows, which completes the proof. ☐

The following theorem, which is a particular claim of Theorem 4.3.7, shows that (1.1.2) holds whenever f therein is absolutely continuous, and is hence an FTC of the second kind.

Theorem 4.3.8 (FTC, version 2)**.** *If* $-\infty < a < b < \infty$ *and* $F : [a,b] \to \mathbb{R}$ *is absolutely continuous, then it is differentiable a.e.,* F' *is integrable on* $[a,b]$ *and*

$$\int_a^b F'(x)\,dx = F(b) - F(a)\,.$$

Proof. Follows from the 'only if' part of Theorem 4.3.7. ☐

The following exercise establishes a connection between Radon-Nikodym derivative, as in Definition 4.1.4, and functional derivative.

Exercise 4.3.4. *Suppose* $F : \mathbb{R} \to \mathbb{R}$ *is a non-decreasing function which is absolutely continuous on* $[a, b]$ *for every compact interval* $[a, b] \subset \mathbb{R}$*. If* μ *is the Riemann-Stieltjes measure of* F*, show that* $\mu \ll \lambda$ *and*

$$\frac{d\mu}{d\lambda} = F' \, .$$

The next result gives an easily checkable sufficient condition under which the claim of Theorem 4.3.8 holds.

Theorem 4.3.9 (FTC, version 3)**.** *If* $F : [a, b] \to \mathbb{R}$ *is differentiable everywhere on* (a, b) *and continuous at* a *and* b *such that*

$$\sup_{x \in (a,b)} |F'(x)| < \infty,$$

then

$$\int_a^b F'(x)\, dx = F(b) - F(a) \, .$$

Proof. It suffices to prove that F is absolutely continuous, because then the proof would follow from Theorem 4.3.8. Denoting

$$K = \sup_{x \in (a,b)} |F'(x)|,$$

it is immediate from the mean value theorem that

$$|F(y) - F(x)| \le K|x - y|, \, a \le x \le y \le b \, .$$

Absolute continuity of F follows immediately from the above, and hence completes the proof. ☐

The following exercise is a very commonly used version of the fundamental theorem of calculus.

Exercise 4.3.5. *Suppose for some* $-\infty \le a < b \le \infty$, $F : (a, b) \to \mathbb{R}$ *is a continuously differentiable function. Assume that*

$$\int_a^b |F'(x)|\, dx < \infty \, .$$

Show that

$$\lim_{x \downarrow a} F(x) \text{ and } \lim_{x \uparrow b} F(x) \text{ exist and are finite.}$$

Denoting by $F(a)$ *and* $F(b)$ *the above limits, respectively, show furthermore that*

$$\int_a^b F'(x)\, dx = F(b) - F(a) \, .$$

Theorem 4.3.10. *Suppose* $(\Omega_1, \mathcal{A}_1, \mu)$ *is a measure space,* $(\Omega_2, \mathcal{A}_2)$ *is a measurable space and* $T : \Omega_1 \to \Omega_2$ *is a measurable function. Define*

$$\nu(A) = \mu\left(T^{-1}A\right), \, A \in \mathcal{A}_2 \, . \tag{4.3.10}$$

Then ν *is a measure on* $(\Omega_2, \mathcal{A}_2)$*.*

Proof. Exercise. □

The following definition is motivated by the above theorem.

Definition 4.3.5. *Under the hypotheses of Theorem 4.3.10,* ν *defined by* (4.3.10) *is the* push-forward measure *of* μ *under* T*, and is denoted by*

$$\nu = \mu \circ T^{-1} \, .$$

Exercise 4.3.6. *If* $(\Omega_1, \mathcal{A}_1)$ *is a measurable space,* $(\Omega_2, \mathcal{A}_2, \mu)$ *is a* σ*-finite measure space and* $T : \Omega_1 \to \Omega_2$ *is a bijection such that* $T^{-1} : \Omega_2 \to \Omega_1$ *is measurable, show that* $T(A) \in \mathcal{A}_2$ *for all* $A \in \mathcal{A}_1$ *and* $\nu : \mathcal{A}_1 \to [0, \infty]$ *defined by*

$$\nu(A) = \mu\big(T(A)\big), \, A \in \mathcal{A}_1, \tag{4.3.11}$$

is a σ*-finite measure on* $(\Omega_2, \mathcal{A}_2)$*.*

The measure ν, defined by (4.3.11), is sometimes called the "pull-back" of μ under T. The pull-back measure is defined whenever μ is a measure and T is a bijection whose inverse is measurable.

Theorem 4.3.11. *Suppose* $-\infty \le a < b \le \infty$ *and* $-\infty \le c < d \le \infty$ *and* $F : (a, b) \to (c, d)$ *is a bijection which is absolutely continuous on* $[\alpha, \beta]$ *for all* $a < \alpha < \beta < b$*. Let* $\mu = \lambda \circ F$*, that is,*

$$\mu(A) = \lambda(F(A)), \, A \in \mathcal{B}\big((a, b)\big) \, .$$

Then μ *is a* σ*-finite measure on* (a, b)*,* $\mu \ll \lambda$ *as measures on* (a, b) *and*

$$\frac{d\mu}{d\lambda}(x) = |F'(x)|, \, x \in (a, b) \, .$$

Proof. Since F is a continuous bijection on (a, b), it is either increasing or decreasing. We prove the result for the case F is decreasing; the proof of the other case is similar.

As F is decreasing, its inverse F^{-1} is also decreasing and hence measurable, Exercise 4.3.6 shows μ is the pull-back of λ under F from (c, d) to (a, b). Since λ is σ-finite, so is μ. Define a measure ν on (a, b) by

$$\nu(A) = \int_A \big(-F'(x)\big) \, dx \, . \tag{4.3.12}$$

In other words, ν is the measure whose Radon-Nikodym derivative with respect to λ is $-F'$. Then for $a < \alpha < \beta < b$,

$$\begin{aligned}\nu\left((\alpha,\beta]\right) &= -\int_\alpha^\beta F'(x)\,dx\\ &= F(\alpha) - F(\beta)\\ &= \lambda\Big(\big[F(\beta), F(\alpha)\big)\Big)\\ &= \mu\big((\alpha,\beta]\big),\end{aligned}$$

the second line following from Theorem 4.3.8 and the list line follows from the observation that since F is decreasing, $F\big((\alpha,\beta]\big) = [F(\beta), F(\alpha))$.

Continuity from below shows that μ and ν agree on

$$\mathcal{S} = \{(\alpha,\beta] \cap (a,b) : a \le \alpha \le \beta \le b\},$$

where $\mathcal{S}$ is a semi-field. Further, μ and ν are σ-finite on $\mathcal{S}$ because if $a < \alpha_n < \beta_n < b$ are such that $\alpha_n \to a$ and $\beta_n \to b$, then

$$(a,b) = \bigcup_{n=1}^\infty (\alpha_n, \beta_n]$$

and

$$\mu\big((\alpha_n,\beta_n]\big) = \nu\big((\alpha_n,\beta_n]\big) = F(\alpha_n) - F(\beta_n) < \infty\,.$$

Theorem 2.5.4 implies μ and ν agree on $\sigma(\mathcal{S})$ which is the same as $\mathcal{B}\big((a,b)\big)$. Hence $\mu \ll \lambda$ and (4.3.12) shows

$$\frac{d\mu}{d\lambda} = -F' = |F'|,$$

which completes the proof. ☐

Exercise 4.3.7. *Suppose* $(\Omega_1, \mathcal{A}_1, \mu)$ *is a measure space,* $(\Omega_2, \mathcal{A}_2)$ *is a measurable space and* $T : \Omega_1 \to \Omega_2$ *is a measurable map. Then for a measurable* $f : \Omega_2 \to \overline{\mathbb{R}}$, *show that*

$$\int_{\Omega_1} f\left(T(\omega)\right)\,\mu(d\omega) = \int_{\Omega_2} f(x)\,\mu\circ T^{-1}(dx), \tag{4.3.13}$$

where the integrals above refer to the notations introduced in (3.3.1) *which is just below Definition 3.3.1, whenever either side of* (4.3.13) *makes sense. This is the so-called change of variables formula for push-forward measures.*

Exercise 4.3.8. *Suppose* $(\Omega, \mathcal{A}, \mu)$ *is a measure space and* $f : \Omega \to [0,\infty]$ *is measurable. Define*

$$\nu(A) = \int_A f\,d\mu,\ A \in \mathcal{A}\,.$$

1. *Show that* ν *is a measure on* $(\Omega, \mathcal{A})$.
2. *Prove that for a measurable* $g : \Omega \to \overline{\mathbb{R}}$,

$$\int g\, d\nu = \int gf\, d\mu, \tag{4.3.14}$$

whenever either side is defined.

The formula (4.3.14) is called the "change of measure formula".

Theorem 4.3.12 (Change of variables for Lebesgue integration)**.** *Suppose* $-\infty \leq a < b \leq \infty$, $-\infty \leq c < d \leq \infty$ *and* $F : (a, b) \to (c, d)$ *is a bijection which is absolutely continuous on* $[\alpha, \beta]$ *for all* $a < \alpha < \beta < b$. *Then for a measurable* $f : (c, d) \to \mathbb{R}$,

$$\int_a^b f \circ F(x) |F'(x)|\, dx = \int_c^d f(y)\, dy, \tag{4.3.15}$$

whenever the integral on either side is defined.

Remark 4.3.1. *1. The above theorem justifies the substitution*

$$y = F(x), \quad dy = |F'(x)|\, dx\,.$$

2. The formula (4.3.15) *becomes incorrect if the modulus sign on* $F'(x)$ *is removed.*

Exercise 4.3.9. *If* $(\Omega_1, \mathcal{A}_1)$ *is a measurable space,* $(\Omega_2, \mathcal{A}_2, \mu)$ *is a measure space and* $T : \Omega_1 \to \Omega_2$ *is a bijection such that* $T^{-1} : \Omega_2 \to \Omega_1$ *is measurable, show that for a measurable* $f : \Omega_2 \to \overline{\mathbb{R}}$,

$$\int_{\Omega_2} f(y)\, \mu(dy) = \int_{\Omega_1} f\big(T(x)\big)\, \mu \circ T(dx)$$

whenever the integral on either side is defined.

Proof of Theorem 4.3.12. Let $\mu = \lambda \circ F$. For a non-negative measurable $f : (c, d) \to \overline{\mathbb{R}}$, Exercise 4.3.9 shows that

$$\begin{aligned} \int_c^d f(y)\, dy &= \int_a^b f \circ F(x)\, \lambda \circ F(dx) \\ &= \int_a^b f \circ F(x)\, \mu(dx) \\ &= \int_a^b f \circ F(x) |F'(x)|\, dx, \end{aligned}$$

(4.3.14) and Theorem 4.3.11 implying the last line. Thus (4.3.15) holds whenever either side of it makes sense. This completes the proof. ☐

We end this subsection with a corollary of Theorem 4.3.12.

Corollary 4.3.1. *Suppose* $U, V \subset \mathbb{R}$ *are open sets and* $F : U \to V$ *is a bijection which is absolutely continuous on* $[\alpha, \beta]$ *whenever* $[\alpha, \beta] \subset U$*. Then for a measurable* $f : V \to \overline{\mathbb{R}}$*,*

$$\int_U f \circ F(x)|F'(x)|\, dx = \int_V f(y)\, dy \tag{4.3.16}$$

whenever either side makes sense.

Once again, (4.3.16) becomes incorrect if the modulus sign on $F'(x)$ is removed.

4.4 Complex Measures

Let $(\Omega, \mathcal{A})$ be a measurable space, fixed for this subsection. The following is in line with Definition 4.1.3.

Definition 4.4.1. *A function* $\mu : \mathcal{A} \to \mathbb{C}$ *is a* complex measure *if*

$$\sum_{n=1}^{\infty} |\mu(A_n)| < \infty$$

and

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

for all disjoint $A_1, A_2, \ldots \in \mathcal{A}$*. The tuple* $(\Omega, \mathcal{A}, \mu)$ *is a* complex measure space*.*

As mention in Remark 4.1.1, the first condition is implied by the second one, though it doesn't hurt to mention the former separately for the sake of emphasis. A signed measure is trivially a complex measure.

Definition 4.4.2. *Given a complex measure space* $(\Omega, \mathcal{A}, \mu)$*, the* total variation $|\mu|(A)$ *of* μ *on* A *for* $A \in \mathcal{A}$ *is defined by*

$$|\mu|(A) = \sup\left\{\sum_{i=1}^{n} |\mu(A_i)| : A_1, \ldots, A_n \in \mathcal{A} \text{ disjoint}, \bigcup_{i=1}^{n} A_i = A,\, n \in \mathbb{N}\right\}.$$

Let μ be a complex measure on $(\Omega, \mathcal{A})$ unless specified otherwise.

Remark 4.4.1. *The total variation* $|\mu|(A)$ *of* μ *on* A *should not be confused with* $|\mu(A)|$*.*

Exercise 4.4.1. *Show that*

$$|\mu(A)| \le |\mu|(A),\ A \in \mathcal{A}.$$

Definition 4.4.3. *If $A \in \mathcal{A}$ is the union of $A_1, A_2, \ldots$ which is a finite or countably infinite collection of disjoint sets in $\mathcal{A}$, then $\{A_1, A_2, \ldots\}$ is a* partition *of A.*

Exercise 4.4.2. *Show that for*

$$|\mu|(A) = \sup \left\{ \sum_{i=1}^{\infty} |\mu(A_i)| : \{ A_1, A_2, \ldots\} \text{ is an infinite partition of } A \right\} .$$

Theorem 4.4.1. *If μ is a signed measure on Ω and (A^+, A^-) is its Hahn decomposition, that is, $A^- = (A^+)^c$ and for all $A \in \mathcal{A}$,*

$$\mu(A) \geq 0 \text{ if } A \subset A^+ \text{ and } \mu(A) \leq 0 \text{ if } A \subset A^-,$$

then

$$|\mu|(A) = \mu(A \cap A^+) - \mu(A \cap A^-),\ A \in \mathcal{A}\,.$$

Proof. Fix $A \in \mathcal{A}$ and let $A_1 = A \cap A^+$ and $A_2 = A \cap A^-$. Then $\{A_1, A_2\}$ is a partition of A and hence

$$|\mu(A)| \geq |\mu(A_1)| + |\mu(A_2)| = \mu(A \cap A^+) - \mu(A \cap A^-)\,.$$

Now let $\{A_1, \ldots, A_n\}$ be any partition of A. Then

$$\begin{aligned} \sum_{i=1}^{n} |\mu(A_i)| &= \sum_{i=1}^{n} |\mu(A_i \cap A^+) + \mu(A_i \cap A^-)| \\ &\leq \sum_{i=1}^{n} \left[|\mu(A_i \cap A^+)| + |\mu(A_i \cap A^-)| \right] \\ &= \sum_{i=1}^{n} \left[\mu(A_i \cap A^+) - \mu(A_i \cap A^-) \right] \\ &= \mu(A \cap A^+) - \mu(A \cap A^-)\,. \end{aligned}$$

As this holds for any partition of A, it follows that

$$|\mu|(A) \leq \mu(A \cap A^+) - \mu(A \cap A^-),$$

which in conjunction with the already proved reverse inequality completes the proof. ☐

The following definition is motivated by the above theorem.

Definition 4.4.4. *For a signed measure μ on $(\Omega, \mathcal{A})$ with Hahn decomposition (A^+, A^-), the* positive and negative parts *of μ, denoted by μ^+ and μ^-, respectively, are defined by*

$$\mu^+(A) = \mu(A \cap A^+) \text{ and } \mu^-(A) = -\mu(A \cap A^-),\ A \in \mathcal{A}\,. \tag{4.4.1}$$

The exercise below follows from Theorem 4.4.1.

Exercise 4.4.3. *Suppose μ is a signed measure on Ω.*

1. *Show that μ^+ and μ^- are finite measures on Ω and are independent of the choice of the Hahn decomposition, that is, if (A_1^+, A_1^-) and (A_2^+, A_2^-) are such that $A_i^+ = (A_i^-)^c$ and*

$$\mu(A) \ge 0 \text{ if } A \subset A_i^+ \text{ and } \mu(A) \le 0 \text{ if } A \subset A_i^-,$$

for $i = 1, 2$, then

$$\mu(A \cap A_1^\pm) = \mu(A \cap A_2^\pm),\ A \in \mathcal{A}.$$

2. *Show that*

$$|\mu|(A) = \mu^+(A) + \mu^-(A),\ A \in \mathcal{A}. \tag{4.4.2}$$

Definition 4.4.5. *For $z = x + \iota y \in \mathbb{C}$ where $x, y \in \mathbb{R}$, the* real and imaginary parts *of z, denoted by $\Re(z)$ and $\Im(z)$, respectively, are defined by*

$$\Re(z) = x \text{ and } \Im(z) = y\,.$$

The notation $\iota = \sqrt{-1}$ is used above and elsewhere in this subsection. It is worth emphasizing that the imaginary part is also a real number, and

$$z = \Re(z) + \iota \Im(z),\ z \in \mathbb{C}\,.$$

Theorem 4.4.2. *The real and complex parts $\Re\mu$ and $\Im\mu$ of the complex measure μ are signed measures. Further for all $A \in \mathcal{A}$,*

$$|\mu|(A) \le |\Re\mu|(A) + |\Im\mu|(A)$$

and

$$|\Re\mu|(A) \vee |\Im\mu|(A) \le |\mu|(A)\,.$$

Proof. That $\Re\mu$ and $\Im\mu$ are signed measures follow trivially since μ is a complex measure. For $A \in \mathcal{A}$ and a partition $\{A_1, \ldots, A_n\}$ of A,

$$\sum_{i=1}^n |\mu(A_i)| \le \sum_{i=1}^n \left(|\Re\mu(A_i)| + |\Im\mu(A_i)|\right) \le |\Re\mu|(A) + |\Im\mu|(A)\,.$$

Taking the supremum over all finite partitions $\{A_1, \ldots, A_n\}$ of A, the first claim follows.

For the second claim, observe that for $A_1, \ldots, A_n$ as above,

$$\sum_{i=1}^n |\mu(A_i)| \ge \sum_{i=1}^n |\Re\mu(A_i)|,$$

which shows by taking a supremum that

$$|\mu|(A) \geq |\Re\mu|(A)\,.$$

This in conjunction with

$$|\mu|(A) \geq |\Im\mu|(A)\,,$$

which follows similarly, proves the second claim. This completes the proof. $\square$

Theorem 4.4.3. *The total variation $|\mu|$ is a finite measure on Ω.*

Proof. Theorem 4.4.2 implies

$$|\mu|(A) \leq |\Re\mu|(A) + |\Im\mu|(A) < \infty,\ A \in \mathcal{A}\,. \tag{4.4.3}$$

To show that $|\mu|$ is a measure, let $A_1, A_2, \ldots \in \mathcal{A}$ be disjoint and set

$$A = \bigcup_{n=1}^{\infty} A_n\,.$$

For a fixed $\varepsilon > 0$, (4.4.3) implies $|\mu|(A_n) - 2^{-n}\varepsilon < |\mu|(A_n)$ and hence there exists a partition $\{A_{n1}, \ldots, A_{nk_n}\}$ of A_n such that

$$\sum_{i=1}^{k_n} |\mu(A_{ni})| \geq |\mu|(A_n) - 2^{-n}\varepsilon\,. \tag{4.4.4}$$

Thus $\{A_{11}, \ldots, A_{1k_1}, A_{21}, \ldots, A_{2k_2}, \ldots\}$ is an infinite partition of A. Exercise 4.4.2 implies that

$$\begin{aligned}
|\mu|(A) &\geq \sum_{n=1}^{\infty}\sum_{i=1}^{k_n} |\mu(A_{ni})| \\
&\geq \sum_{n=1}^{\infty} \big(|\mu|(A_n) - 2^{-n}\varepsilon\big) \\
&= -\varepsilon + \sum_{n=1}^{\infty} |\mu|(A_n),
\end{aligned}$$

(4.4.4) implying the inequality in the second line. Since ε is arbitrary, it follows that

$$|\mu|(A) \geq \sum_{n=1}^{\infty} |\mu|(A_n)\,.$$

For the reverse inequality, let $\{E_1, \ldots, E_k\}$ be a partition of A. Then

$$E_i = \bigcup_{n=1}^{\infty} (E_i \cap A_n),\ i = 1, \ldots, k,$$

and the sets on the right hand side are disjoint. Hence

$$\begin{aligned}\sum_{i=1}^{k}|\mu(E_i)| &= \sum_{i=1}^{k}\left|\sum_{n=1}^{\infty}\mu(E_i\cap A_n)\right|\\ &\le \sum_{i=1}^{k}\sum_{n=1}^{\infty}|\mu(E_i\cap A_n)|\\ &= \sum_{n=1}^{\infty}\sum_{i=1}^{k}|\mu(E_i\cap A_n)|\\ &\le \sum_{n=1}^{\infty}|\mu|(A_n),\end{aligned}$$

the last line following from the observation that $A_n\cap E_1,\ldots,A_n\cap E_k$ is a partition of A_n for all $n\in\mathbb{N}$. Taking supremum over all finite partitions $\{E_1,\ldots,E_k\}$ of A, we get

$$|\mu|(A)\le\sum_{n=1}^{\infty}|\mu|(A_n)\,.$$

This along with the already proven reverse inequality shows $|\mu|$ is a measure. Putting $A=\Omega$ in (4.4.3) shows that $|\mu|$ is a finite measure and hence the proof follows. □

Definition 4.4.6. *The* Borel σ-field on the complex plane*, denoted by* $\mathcal{B}(\mathbb{C})$*, is the* σ*-field on* $\mathbb{C}$ *generated by open sets.*

Exercise 4.4.4. *1. Show that*

$$\mathcal{B}(\mathbb{C})=\sigma\left(\{R_{abcd}:a,b,c,d\in\mathbb{R},\ a<b,\ c<d\}\right)$$

where

$$R_{abcd}=\{z\in\mathbb{C}:a<\Re(z)<b,\ c<\Im(z)<d\},\ a,b,c,d\in\mathbb{R},\ a<b,\ c<d\,.$$

2. Using 1. or otherwise, show that $f:\Omega\to\mathbb{C}$ *is* $\mathcal{B}(\mathbb{C})$*-measurable if and only if* $\Re f$ *and* $\Im f$ *are Borel functions.*

Henceforth ν is a measure on $(\Omega,\mathcal{A})$ and μ continues to be a complex measure on the same space, unless specified otherwise.

Definition 4.4.7. *If* $f:\Omega\to\mathbb{C}$ *is such that* $\Re f,\Im f$ *are in* $L^1(\Omega,\nu)$*, then* f *is said to be* integrable with respect to ν *and we define*

$$\int f\,d\nu=\int\Re f\,d\nu+i\int\Im f\,d\nu\,.$$

Exercise 4.4.5. *For a measurable* $f : \Omega \to \mathbb{C}$*, show that* $|f|$ *is Borel and* f *is integrable with respect to* ν *if and only if* $|f|$ *is* ν*-integrable.*

Exercise 4.4.6. *For* ν*-integrable* $f, g : \Omega \to \mathbb{C}$ *and* $z_1, z_2 \in \mathbb{C}$*, show that* $z_1 f + z_2 g$ *is also* ν*-integrable and*

$$\int (z_1 f + z_2 g)\, d\nu = z_1 \int f\, d\nu + z_2 \int g\, d\nu\,.$$

Theorem 4.4.4. *For a* ν*-integrable* $f : \Omega \to \mathbb{C}$*,*

$$\left|\int f\, d\nu\right| \le \int |f|\, d\nu\,.$$

Proof. Let $z = \int f\, d\nu$. If $z = 0$, the claimed inequality follows trivially. Assume without loss of generality that $z \neq 0$. Then

$$1 = z^{-1} \int f\, d\nu = \int z^{-1} f\, d\nu,$$

the right equality following from Exercise 4.4.6. Thus $\int z^{-1} f\, d\nu$ is a real number. It follows from the definition that

$$\int \Im(z^{-1} f)\, d\nu = \Im\left(\int z^{-1} f\, d\nu\right) = 0\,.$$

Therefore

$$\begin{aligned} 1 &= \int \Re(z^{-1} f)\, d\nu \\ &= \left|\int \Re(z^{-1} f)\, d\nu\right| \\ &\le \int |\Re(z^{-1} f)|\, d\nu \\ &\le \int |z^{-1} f|\, d\nu \\ &= |z|^{-1} \int |f|\, d\nu, \end{aligned}$$

Hence

$$|z| \le \int |f|\, d\nu,$$

which is precisely the claimed inequality. Hence the proof. ☐

Theorem 4.4.5. *If* $f : \Omega \to [0, \infty]$ *is measurable, then*

$$\int f\, d(\Re\mu)^{\pm} \le \int f\, d|\mu|,$$

$$\int f\,d(\Im\mu)^{\pm} \le \int f\,d|\mu|$$

and

$$\int f\,d|\mu| \le \int f\,d(\Re\mu)^{+} + \int f\,d(\Re\mu)^{-} + \int f\,d(\Im\mu)^{+} + \int f\,d(\Im\mu)^{-}\,.$$

Proof. If $f = \mathbf{1}_A$ for some $A \in \mathcal{A}$, then

$$\int f\,d(\Re\mu)^{+} = (\Re\mu)^{+}(A) \le |\Re\mu|(A) \le |\mu|(A) = \int f\,d|\mu|,$$

(4.4.2) implying the first inequality and the second one following from the second claim of Theorem 4.4.2. Routine arguments show the claim for a non-negative simple function f and the monotone convergence theorem shows that

$$\int f\,d(\Re\mu)^{+} \le \int f\,d|\mu|\,.$$

A similar argument works for $(\Re\mu)^{-}$ and $(\Im\mu)^{\pm}$ and hence proves the first two claims.

The third claim also follows from (4.4.2) and the first inequality of Theorem 4.4.2 in a similar way. This completes the proof. ☐

Definition 4.4.8. *A measurable* $f : \Omega \to \mathbb{C}$ *is* μ-integrable *if*

$$\int |f|\,d|\mu| < \infty,$$

and for such f*, its* integral with respect to μ *is defined by*

$$\int f\,d\mu = \int f\,d\,(\Re\mu)^{+} - \int f\,d\,(\Re\mu)^{-} + \iota\int f\,d\,(\Im\mu)^{+} - \iota\int f\,d\,(\Im\mu)^{-}\,;$$

each integral on the right hand side is finite due to Theorem 4.4.5.

The notion of absolute continuity of a complex measure with respect to a measure is as in Definition 4.2.3 and is not being stated explicitly.

Theorem 4.4.6 (Radon-Nikodym theorem for complex measures)**.** *Assume* $\mu \ll \nu$*, that is,* $\nu(A) = 0$ *implies that* $\mu(A) = 0$ *and that* ν *is* σ*-finite. Then there exists a* ν*-integrable* $f : \Omega \to \mathbb{C}$ *such that*

$$\mu(A) = \int_A f\,d\nu,\ A \in \mathcal{A}\,.$$

Proof. Observe that $\mu \ll \nu$ implies that $\Re\mu \ll \nu$ and $\Im\mu \ll \nu$. Since ν is σ-finite, the proof follows by applying Theorem 4.2.5 to $(\Re\mu, \nu)$ and $(\Im\mu, \nu)$. ☐

Theorem 4.4.7. *The exists $h : \Omega \to \mathbb{C}$ which is $|\mu|$-integrable such that for all A,*

$$\mu(A) = \int_A h \, d|\mu| \, . \tag{4.4.5}$$

Proof. Exercise 4.4.1 shows that $|\mu|(A) = 0$ implies

$$|\mu(A)| \le |\mu|(A) = 0$$

and hence $\mu \ll |\mu|$. Since $|\mu|$ is a finite measure, Theorem 4.4.6 completes the proof. ☐

Theorem 4.4.8. *If h is as in* (4.4.5)*, then*

$$|h| = 1, \ |\mu|\text{-a.e.}$$

The proof of the above theorem uses the following lemma.

Lemma 4.4.1. *Suppose ν is a finite measure on* $(\Omega, \mathcal{A})$ *and $f : \Omega \to \mathbb{C}$ is ν-integrable. If $S \subset \mathbb{C}$ is a closed set such that*

$$\frac{1}{\nu(E)} \int_E f \, d\nu \in S$$

for all $E \in \mathcal{A}$ with $\nu(E) > 0$*, then $f \in S$ ν-a.e.*

Proof. Since S^c is an open set, as S is closed, there exist $z_1, z_2, \ldots \in S^c$ and $r_1, r_2, \ldots > 0$ such that

$$S^c = \bigcup_{n=1}^{\infty} \{z \in \mathbb{C} : |z - z_n| \le r_n\} \, .$$

Thus it suffices to show that

$$\nu\left(f^{-1}\{z \in \mathbb{C} : |z - z_0| \le r\}\right) = 0$$

whenever z_0, r are such that $\{z \in \mathbb{C} : |z - z_0| \le r\} \subset S^c$.

Let z_0, r be as above. Denote $E = f^{-1}\{z \in \mathbb{C} : |z - z_0| \le r\}$. For showing $\nu(E) = 0$, assume the contrary, that is, $\nu(E) > 0$. Then

$$\begin{aligned}
\left| \frac{1}{\nu(E)} \int_E f \, d\nu - z_0 \right| &= \left| \frac{1}{\nu(E)} \int_E (f - z_0) \, d\nu \right| \\
&\le \frac{1}{\nu(E)} \int_E |f - z_0| \, d\nu \\
&\le \frac{1}{\nu(E)} \int_E r \, d\nu \\
&= r,
\end{aligned}$$

Theorem 4.4.4 implying the inequality in the second line and that in the penultimate line following from the definition of E. Thus

$$\frac{1}{\nu(E)} \int_E f \, d\nu \in \{z \in \mathbb{C} : |z - z_0| \le r\} \subset S^c,$$

which contradicts the hypothesis of the lemma. Hence $\nu(E) = 0$, from which the proof follows. $\square$

Proof of Theorem 4.4.8. Let h be as in (4.4.5) and $S = \{z \in \mathbb{C} : |z| \le 1\}$. Then for all $E \in \mathcal{A}$ with $|\mu|(E) > 0$,

$$\left| \frac{1}{|\mu|(E)} \int h \, d\mu \right| = \frac{|\mu(E)|}{|\mu|(E)} \le 1,$$

the inequality following from Exercise 4.4.1. In other words,

$$\frac{1}{|\mu|(E)} \int h \, d\mu \in S \text{ whenever } |\mu|(E) > 0 \,.$$

Lemma 4.4.1 shows $h \in S$ $|\mu|$-a.e. That is,

$$|h| \le 1, \ |\mu|\text{-a.e.} \tag{4.4.6}$$

Let $\{A_1, A_2, \dots, A_n\}$ be a partition of $A = [|h| \le 1 - \varepsilon]$ for some $\varepsilon > 0$. Then

$$\sum_{i=1}^n |\mu(A_i)| = \sum_{i=1}^n \left| \int_{A_i} h \, d|\mu| \right| \le \sum_{i=1}^n \int_{A_i} |h| \, d|\mu| \le (1 - \varepsilon)|\mu|(A) \,.$$

Since this holds for every partition,

$$|\mu|(A) \le (1 - \varepsilon)|\mu|(A) \,.$$

Hence $|\mu|(A) = 0$, that is, $\mu(|h| \le 1 - \varepsilon) = 0$ for all $0 < \varepsilon < 1$. Recalling that

$$[|h| < 1] = \bigcup_{k=1}^\infty [|h| \le 1 - 1/k],$$

it follows that $|h| \ge 1$ $|\mu|$-a.e. The proof follows by combining this with (4.4.6). $\square$

Theorem 4.4.9 (Change of measure formula for complex measures)**.** *Let ν be a measure on Ω and $f : \Omega \to \mathbb{C}$ be ν-integrable. Define*

$$\mu(A) = \int_A f \, d\nu, \ A \in \mathcal{A} \,.$$

Then μ is a complex measure on Ω. A measurable $g : \Omega \to \mathbb{C}$ is μ-integrable if and only if $|fg| \in L^1(\Omega, \nu)$ and for such g,

$$\int g \, d\mu = \int fg \, d\nu \,.$$

Proof. Let $f_1 = (\Re f)^+$, $f_2 = (\Re f)^-$, $f_3 = (\Im f)^+$ and $f_4 = (\Im f)^-$ where $x^+ = x \vee 0$ and $x^- = (-x) \vee 0$ for $x \in \mathbb{R}$, as usual. Recall that

$$|f_1| \vee \ldots \vee |f_4| \le |f|,$$

and hence $f_1, \ldots, f_4$ are ν-integrable. Define

$$\mu_i(A) = \int_A f_i \, d\nu, \; i = 1, 2, 3, 4 \, .$$

Thus $\mu_1, \ldots, \mu_4$ are finite measures on Ω. Since

$$\mu = \mu_1 - \mu_2 + \iota\mu_3 - \iota\mu_4,$$

it follows that μ is a complex measure on Ω.

A trivial observation is that

$$\Re\mu(A) = \int_A \Re f \, d\nu, \; A \in \mathcal{A} \, .$$

Hence $([\Re f \ge 0], [\Re f < 0])$ is a Hahn decomposition for $\Re\mu$. Thus for all $A \in \mathcal{A}$,

$$(\Re\mu)^+(A) = \mu_1(A), \tag{4.4.7}$$

$$(\Re\mu)^-(A) = \mu_2(A) \, . \tag{4.4.8}$$

Similarly, it can be shown that

$$(\Im\mu)^+(A) = \mu_3(A), \tag{4.4.9}$$

$$(\Im\mu)^-(A) = \mu_4(A) \, . \tag{4.4.10}$$

Let $g : \Omega \to \mathbb{C}$ be measurable. Suppose that g is μ-integrable, that is,

$$\int |g| \, d|\mu| < \infty \, .$$

The first claim of Theorem 4.4.5 implies that

$$\begin{aligned} \int |g| \, d|\mu| &\ge \int |g| \, d(\Re\mu)^+ \\ &= \int |g| \, d\mu_1 \\ &= \int |g| f_1 \, d\nu, \end{aligned}$$

(4.4.7) implying the second line and the last line following from the definition of μ_1 and (4.3.14). Similarly, (4.4.8)-(4.4.10) imply

$$\int |g| f_i \, d\nu \le \int |g| \, d|\mu|, \; i = 2, 3, 4 \, .$$

Thus

$$\int |fg|\,d\nu \le \sum_{i=1}^4 \int |g|f_i d\nu \le 4\int |g|\,d|\mu| < \infty\,.$$

Conversely, suppose that

$$\int |fg|\,d\nu < \infty\,.$$

The third claim of Theorem 4.4.5 and (4.4.7)-(4.4.10) imply that

$$\begin{aligned}\int |g|\,d|\mu| &\le \sum_{i=1}^4 \int |g|\,d\mu_i\\ &= \sum_{i=1}^4 \int |g|f_i\,d\nu\\ &\le 2\int |fg|\,d\nu\\ &< \infty,\end{aligned}$$

(4.3.14) and the definition of $\mu_1, \ldots, \mu_4$ implying the second line and the inequality in the penultimate line following from the observation

$$\sum_{i=1}^4 f_i = |\Re f| + |\Im f| \le 2|f|\,.$$

This proves the "if and only if" claim.

Finally let g be μ-integrable. It follows from (4.3.14) that

$$\int \Re g\,d\mu_i = \int f_i \Re g\,d\nu,\ i = 1, \ldots, 4\,.$$

A similar argument works for $\Im g$, which in conjunction with the above shows

$$\int g\,d\mu_i = \int f_i g\,d\nu,\ i = 1, \ldots, 4\,. \tag{4.4.11}$$

Recalling (4.4.7)-(4.4.10) and Definition 4.4.8, we get

$$\begin{aligned}\int g\,d\mu &= \int g\,d\mu_1 - \int g\,d\mu_2 + \iota\int g\,d\mu_3 - \iota\int g\,d\mu_4\\ &= \int f_1 g\,d\nu - \int f_2 g\,d\nu + \iota\int f_3 g\,d\nu - \iota\int f_4 g\,d\nu\\ &= \int fg\,d\nu,\end{aligned}$$

(4.4.11) implying the second line. This completes the proof. □

The following result is the goal of this subsection.

Theorem 4.4.10. *If $f : \Omega \to \mathbb{C}$ is μ-integrable, then*

$$\left|\int f\, d\mu\right| \le \int |f|\, d|\mu|\,.$$

Proof. Theorem 4.4.7 implies the existence of h satisfying (4.4.5). Thus for a μ-integrable $f : \Omega \to \mathbb{C}$, Theorem 4.4.9 shows that

$$\begin{aligned}\left|\int f\, d\mu\right| &= \left|\int fh\, d|\mu|\right| \\ &\le \int |fh|\, d|\mu| \\ &= \int |f|\, d|\mu|,\end{aligned}$$

the second line and the last line following from Theorems 4.4.4 and 4.4.8, respectively. This completes the proof. □

5 Product measures

5.1 The Theorems of Tonelli and Fubini

Suppose $(\Omega_1, \mathcal{A}_1, \mu_1)$ and $(\Omega_2, \mathcal{A}_2, \mu_2)$ are finite measure spaces. Define

$$\Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2) : \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}\,, \tag{5.1.1}$$

that is, $\Omega_1 \times \Omega_2$ is the usual Cartesian product of Ω_1 and Ω_2, and let

$$\mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma\left(\{A_1 \times A_2 : A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\}\right)\,. \tag{5.1.2}$$

For all $E \subset \Omega_1 \times \Omega_2$, define

$$\begin{aligned}E_{\omega_1} &= \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in E\}\,, \omega_1 \in \Omega_1, \\ E^{\omega_2} &= \{\omega_1 \in \Omega_1 : (\omega_1, \omega_2) \in E\}\,, \omega_2 \in \Omega_2\,.\end{aligned}$$

The following result is the first step towards constructing product measures.

Theorem 5.1.1. *For all fixed $E \in \mathcal{A}_1 \otimes \mathcal{A}_2$, the following holds:*

1. *$E_{\omega_1} \in \mathcal{A}_2$ for all $\omega_1 \in \Omega_1$,*
2. *$\omega_1 \mapsto \mu_2(E_{\omega_1})$ is an $\mathcal{A}_1$-measurable function on Ω_1 and*
3. *μ defined on $\mathcal{A}_1 \otimes \mathcal{A}_2$ by*

   $$\mu(E) = \int_{\Omega_1} \mu_2(E_{\omega_1})\, \mu_1(d\omega_1),\ E \in \mathcal{A}_1 \otimes \mathcal{A}_2,$$

   is a measure on $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$.

Similarly, the following holds for all $E \in \mathcal{A}_1 \otimes \mathcal{A}_2$*:*

4. $E^{\omega_2} \in \mathcal{A}_1$ *for all* $\omega_2 \in \Omega_2$*,*
5. $\omega_2 \mapsto \mu_1(E^{\omega_2})$ *is an* $\mathcal{A}_2$*-measurable function on* Ω_2 *and*
6. μ' *defined on* $\mathcal{A}_1 \otimes \mathcal{A}_2$ *by*

$$\mu'(E) = \int_{\Omega_2} \mu_1(E^{\omega_2})\, \mu_2(d\omega_2),\ E \in \mathcal{A}_1 \otimes \mathcal{A}_2,$$

is a measure on $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$*.*

Finally,

$$\mu(E) = \mu'(E) \text{ for all } E \in \mathcal{A}_1 \otimes \mathcal{A}_2 \,. \tag{5.1.3}$$

Proof. The proof follows by routine verifications, as sketched below. Let

$$\mathcal{G} = \{A \subset \Omega_1 \times \Omega_2 :\ \text{each of 1,2, 4 and 5 above holds for } A\}\,.$$

Define

$$\mathcal{S} = \{A_1 \times A_2 :\ A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\},$$

which is a semi-field because it is trivially closed under finite intersections and for $A_1 \in \mathcal{A}_1$ and $A_2 \in \mathcal{A}_2$,

$$(A_1 \times A_2)^c = (A_1 \times A_2^c) \cup (A_1^c \times \Omega_2)\ .$$

It is immediate that $\mathcal{S} \subset \mathcal{G}$ because for $A_1 \times A_2 \in \mathcal{S}$,

$$(A_1 \times A_2)_{\omega_1} = \begin{cases} A_2, & \omega_1 \in A_1, \\ \emptyset, & \text{else,} \end{cases}$$

and hence

$$\mu_2\left((A_1 \times A_2)_{\omega_1}\right) = \mu_2(A_2)\mathbf{1}_{A_1}(\omega_1) \tag{5.1.4}$$

is an $\mathcal{A}_1$-measurable function of ω_1, showing 1 and 2 hold; 4 and 5 hold for $A_1 \times A_2$ by similar arguments. Routine arguments show that the field generated by $\mathcal{S}$ is contained in $\mathcal{G}$. Finally $\mathcal{G}$ can be shown to be a monotone class with the help of DCT, using the fact that μ_1, μ_2 are finite measures, by standard arguments. The monotone class theorem shows $\mathcal{G} \supset \sigma(\mathcal{S}) = \mathcal{A}_1 \otimes \mathcal{A}_2$, that is, 1, 2, 4 and 5 hold for all $E \in \mathcal{A}_1 \otimes \mathcal{A}_2$.

Once the RHS of 3 is defined, due to 2, that μ is a measure is immediate. Similarly, 6 also follows. For (5.1.3), observe that (5.1.4) implies for all $A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2$,

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

Similarly,

$$\mu'(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2),\ A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\,.$$

Thus μ and μ' are finite measures which agree on the semi-field $\mathcal{S}$. Hence (5.1.3) follows. ☐

The jump from finite to σ-finite is easy and is left as an exercise.

Theorem 5.1.2. *The claims of Theorem 5.1.1 hold when* μ_1 *and* μ_2 *are* σ*-finite measures.*

Proof. Exercise. ☐

Now we are in a position to define the product of two σ-finite measure spaces.

Definition 5.1.1. *Suppose* $(\Omega_1, \mathcal{A}_1, \mu_1)$ *and* $(\Omega_2, \mathcal{A}_2, \mu_2)$ *are* σ*-finite measure spaces. Let* $\Omega = \Omega_1 \times \Omega_2$ *and* $\mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2$ *be as in* (5.1.1) *and* (5.1.2)*, respectively. The unique* σ*-finite measure* μ *on* $(\Omega, \mathcal{A})$ *satisfying*

$$\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2) \text{ for all } A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2$$

is the product measure *of* μ_1 *and* μ_2*. The measure space* $(\Omega, \mathcal{A}, \mu)$ *is the* product of the measure spaces $(\Omega_1, \mathcal{A}_1, \mu_1)$ *and* $(\Omega_2, \mathcal{A}_2, \mu_2)$*, which is written as*

$$(\Omega_1, \mathcal{A}_1, \mu_1) \otimes (\Omega_2, \mathcal{A}_2, \mu_2) = (\Omega, \mathcal{A}, \mu)\,.$$

Usually we also write $\mu = \mu_1 \otimes \mu_2$*.*

Throughout this subsection, $(\Omega_i, \mathcal{A}_i, \mu_i)$ is a σ-finite measure space for $i = 1, 2$.

Theorem 5.1.3 (Tonelli)**.** *Suppose* $f : \Omega_1 \times \Omega_2 \to [0, \infty]$ *is* $\mathcal{A}_1 \otimes \mathcal{A}_2$*-measurable. Then*

1. *for all fixed* $\omega_1 \in \Omega_1$*,* $f(\omega_1, \cdot)$ *is an* $\mathcal{A}_2$*-measurable function from* Ω_2 *to* $[0, \infty]$*,*
2. $\omega_1 \mapsto \int_{\Omega_2} f(\omega_1, \omega_2)\mu_2(d\omega_2)$ *is an* $\mathcal{A}_1$*-measurable function of* ω_1 *on* Ω_1*,*
3. *for all fixed* $\omega_2 \in \Omega_2$*,* $f(\cdot, \omega_2)$ *is an* $\mathcal{A}_1$*-measurable function from* Ω_1 *to* $[0, \infty]$*,*
4. $\omega_2 \mapsto \int_{\Omega_1} f(\omega_1, \omega_2)\mu_1(d\omega_1)$ *is an* $\mathcal{A}_2$*-measurable function of* ω_2 *on* Ω_2*,*
5. *and*

$$\begin{aligned}\int_{\Omega_1}\int_{\Omega_2} f(\omega_1, \omega_2)\,\mu_2(d\omega_2)\,\mu_1(d\omega_1) &= \int_{\Omega_1 \times \Omega_2} f\, d\,(\mu_1 \otimes \mu_2) \\ &= \int_{\Omega_2}\int_{\Omega_1} f(\omega_1, \omega_2)\,\mu_1(d\omega_1)\,\mu_2(d\omega_2)\end{aligned}$$

where

$$\int_{\Omega_1}\int_{\Omega_2} f(\omega_1, \omega_2)\,\mu_2(d\omega_2)\,\mu_1(d\omega_1)$$

means

$$\int_{\Omega_1}\left(\int_{\Omega_2} f(\omega_1, \omega_2)\,\mu_2(d\omega_2)\right)\,\mu_1(d\omega_1)$$

by convention.

Proof. For $f = \mathbf{1}_E$ where $E \in \mathcal{A}_1 \otimes \mathcal{A}_2$, 1, 2, 4, 5 and (5.1.3) of Theorem 5.1.1, which hold for σ-finite measure spaces by Theorem 5.1.2, prove 1-5 above. Routine arguments via simple functions and MCT complete the proof. □

Theorem 5.1.4 (Fubini)**.** *Suppose* $(\Omega, \mathcal{A}, \mu) = (\Omega_1, \mathcal{A}_1, \mu_1) \otimes (\Omega_2, \mathcal{A}_2, \mu_2)$ *and*

$$f \in L^1(\Omega, \mathcal{A}, \mu)\ .$$

Then the following holds.

1. *For all* $\omega_1 \in \Omega_1$, $f(\omega_1, \cdot)$ *is an* $\mathcal{A}_2$*-measurable function form* Ω_2 *to* $\overline{\mathbb{R}}$.
2. *For* (μ_1) *almost all* $\omega_1 \in \Omega_1$,
$$\int_{\Omega_2} |f(\omega_1, \omega_2)|\, \mu_2(d\omega_2) < \infty\,.$$
3. *If* $g : \Omega_1 \to \mathbb{R}$ *is defined by*
$$g(\omega_1) = \begin{cases} \int_{\Omega_2} f(\omega_1, \omega_2)\, \mu_2(d\omega_2), & \text{if } \int_{\Omega_2} |f(\omega_1, \omega_2)|\, \mu_2(d\omega_2) < \infty, \\ 0, & \text{else,} \end{cases}$$
then g *is measurable,* μ_1*-integrable and*
$$\int_{\Omega_1} g\, d\mu_1 = \int_{\Omega} f\, d\mu\,.$$
4. *For all* $\omega_2 \in \Omega_2$, $f(\cdot, \omega_2)$ *is an* $\mathcal{A}_1$*-measurable function form* Ω_1 *to* $\overline{\mathbb{R}}$.
5. *For* (μ_2) *almost all* $\omega_2 \in \Omega_2$,
$$\int_{\Omega_1} |f(\omega_1, \omega_2)|\, mu_1(d\omega_1) < \infty\,.$$
6. *If* $h : \Omega_2 \to \mathbb{R}$ *is defined by*
$$h(\omega_2) = \begin{cases} \int_{\Omega_1} f(\omega_1, \omega_2)\, \mu_1(d\omega_1), & \text{if } \int_{\Omega_1} |f(\omega_1, \omega_2)|\, \mu_1(d\omega_1) < \infty, \\ 0, & \text{else,} \end{cases}$$
then h *is measurable,* μ_2*-integrable and*
$$\int_{\Omega_2} h\, d\mu_2 = \int_{\Omega} f\, d\mu\,.$$

Proof. Follows by applying Theorem 5.1.3 to f^+ and f^-. □

The following corollary of Theorems 5.1.3 and 5.1.4 is often referred to jointly as the Tonelli-Fubini theorem.

Corollary 5.1.1. *For a measurable* $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$*, it holds that*

$$\int_{\Omega_1} \int_{\Omega_2} f(\omega_1, \omega_2)\, \mu_2(d\omega_2)\, \mu_1(d\omega_1) = \int_{\Omega_2} \int_{\Omega_1} f(\omega_1, \omega_2)\, \mu_1(d\omega_1)\, \mu_2(d\omega_2),$$

whenever either $f \geq 0$ *or*

$$\int_{\Omega_1 \times \Omega_2} |f|\, d(\mu_1 \otimes \mu_2) < \infty\,.$$

Example 5.1.1. *Letting* $(\Omega_i, \mathcal{A}_i, \mu_i) = (\mathbb{N}, 2^{\mathbb{N}}, \nu)$ *for* $i = 1, 2$*, where* ν *is the counting measure on* $\mathbb{N}$*, Corollary 5.1.1 implies that for a double array* $(a_{ij} : i, j \in \mathbb{N}) \subset \overline{\mathbb{R}}$*,*

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij} \tag{5.1.5}$$

holds if either

$$a_{ij} \geq 0 \text{ for all } i, j \in \mathbb{N}, \tag{5.1.6}$$

or

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |a_{ij}| < \infty\,. \tag{5.1.7}$$

In view of the above example, it can be asked if (5.1.7) always holds regardless of whether or not (5.1.5) or (5.1.6) is satisfied. The following counter-example answers this question in the negative.

Example 5.1.2. *For all* $i, j \in \mathbb{N}$*, define*

$$a_{ij} = \begin{cases} -1, & j = i, \\ 1, & j = i + 1, \\ 0, & \text{else.} \end{cases}$$

Then

$$\sum_{j=1}^{\infty} a_{ij} = 0 \text{ for all } i = 1, 2, \ldots$$

and hence

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} = 0\,.$$

On the other hand,

$$\sum_{i=1}^{\infty} a_{i1} = -1$$

because $a_{11} = -1$ *and* $a_{i1} = 0$ *for all* $i = 2, 3, \ldots$*; further,*

$$\sum_{i=1}^{\infty} a_{ij} = a_{j-1\,j} + a_{jj} = 1 - 1 = 0,\ j \geq 2\,.$$

Thus

$$\sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij} = -1 \neq 0 = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij},$$

showing that (5.1.5) *does not hold. Neither* (5.1.6) *nor* (5.1.7) *is satisfied in this case and hence the failure of* (5.1.5) *does not lead to any contradiction.*

Along the lines of Definition 5.1.1, the product of σ-finite product measures $(\Omega_1, \mathcal{A}_1, \mu_1), \ldots, (\Omega_n, \mathcal{A}_n, \mu_n)$ can be defined. The usual notation for the same is

$$\bigotimes_{i=1}^{n} (\Omega_i, \mathcal{A}_i, \mu_i) = \left(\prod_{i=1}^{n} \Omega_i, \bigotimes_{i=1}^{n} \mathcal{A}_i, \bigotimes_{i=1}^{n} \mu_i \right),$$

where $\prod_{i=1}^{n} \Omega_i = \Omega_1 \times \ldots \times \Omega_n$.

5.2 The Lebesgue measure in higher dimensions

This subsection is devoted to defining the Lebesgue measure on $\mathbb{R}^d$ and studying its properties. The dimension $d \geq 1$ is fixed throughout.

Definition 5.2.1. *The* Lebesgue measure on $\mathbb{R}^d$*, usually denoted by* λ_d*, is the d-fold product of the Lebesgue measure on* $\mathbb{R}$*, that is,*

$$\bigotimes_{i=1}^{d} \left(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda\right) = \left(\mathbb{R}^d, \bigotimes_{i=1}^{d} \mathcal{B}(\mathbb{R}), \lambda_d \right),$$

where λ *is the Lebesgue measure on* $\mathbb{R}$*. For a measurable* $f : \mathbb{R}^d \to \overline{\mathbb{R}}$*, its* integral with respect to λ_d *is usually denoted by*

$$\int f\, d\lambda_d,\ \int_{\mathbb{R}^d} f(x)\, dx,\ \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(x_1, \ldots, x_d)\, dx_1 \ldots dx_d \ \textit{etc.}$$

The subscript 'd' in λ_d will be suppressed whenever the dimension of the Euclidean space, on which we are working, is obvious from the context.

Definition 5.2.2. *The* Borel σ-field on $\mathbb{R}^d$*, denoted by* $\mathcal{B}(\mathbb{R}^d)$*, is the* σ*-field generated by open sets in* $\mathbb{R}^d$*.*

Exercise 5.2.1. *Show that*

$$\bigotimes_{i=1}^{d} \mathcal{B}(\mathbb{R}) = \mathcal{B}\left(\mathbb{R}^d\right).$$

Let us start with seeing an interesting application of Corollary 5.1.1 in the context of λ_2.

Example 5.2.1. *We wish to calculate*

$$I = \int_{\mathbb{R}} e^{-x^2}\, dx\, .$$

Corollary 5.1.1 implies that

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)}\, dx\, dy &= \int_{-\infty}^{\infty} e^{-x^2} \int_{-\infty}^{\infty} e^{-y^2}\, dy\, dx, \\ &= I^2\, .\end{aligned}$$

For a fixed $x \in \mathbb{R} \setminus \{0\}$, *put* $y = xz$, $dy = |x|\, dz$ *using Theorem 4.3.12 to get*

$$\int_{-\infty}^{\infty} e^{-y^2}\, dy = |x| \int_{-\infty}^{\infty} e^{-z^2x^2}\, dz\, .$$

Thus

$$\begin{aligned}I^2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x| e^{-x^2(1+z^2)}\, dz\, dx \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x| e^{-x^2(1+z^2)}\, dx\, dz,\end{aligned}$$

the second line following once again from Corollary 5.1.1. For a fixed $z \in \mathbb{R}$,

$$\begin{aligned}\int_{-\infty}^{\infty} |x| e^{-x^2(1+z^2)}\, dx &= 2\int_{0}^{\infty} x e^{-x^2(1+z^2)}\, dx \\ (\textit{Theorem 4.3.12: } y = x^2,\ dy = 2x\, dx) &= \int_0^{\infty} e^{-y(1+z^2)}\, dz \\ (\textit{Exercise 4.3.5}) &= \left[-\frac{e^{-y(1+z^2)}}{1+z^2}\right]_{y=0}^{y=\infty} \\ &= \frac{1}{1+z^2}\, .\end{aligned}$$

Therefore

$$\begin{aligned}I^2 &= \int_{-\infty}^{\infty} \frac{1}{1+z^2}\, dz \\ (\textit{Exercise 4.3.5}) &= \left[\tan^{-1} z\right]_{-\infty}^{\infty} \\ &= \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi\, .\end{aligned}$$

Since $I \geq 0$, *we get*

$$I = \sqrt{\pi}\, .$$

that is,

$$\int_{-\infty}^{\infty} e^{-x^2}\, dx = \sqrt{\pi}\,.$$

Theorem 4.3.12 shows that

$$\int_{-\infty}^{\infty} e^{-y^2/2}\, dy = \sqrt{2\pi}\,.$$

In other words,

$$\int_{\mathbb{R}} \varphi(z)\, dz = 1,$$

where $\varphi : \mathbb{R} \to \mathbb{R}$ *is defined by*

$$\varphi(z) = (2\pi)^{-1/2} e^{-z^2/2},\ z \in \mathbb{R}\,.$$

The function φ *is of central importance in probability theory and is called the "Gaussian or normal density" therein.*

Let $\mathcal{H}$ be the collection of all half-open (left open, right closed) bounded rectangles, that is,

$$\mathcal{H} = \{(a_1, b_1] \times \ldots \times (a_d, b_d] : -\infty < a_i < b_i < \infty \text{ for } i = 1, \ldots, d\}\ .$$

Exercise 5.2.2. *Check that*

$$\sigma(\mathcal{H}) = \mathcal{B}(\mathbb{R}^d)\,.$$

For all $R \in \mathcal{H}$, R^0 and $\overline{R}$ denote the interior and closure of R, respectively.

Exercise 5.2.3. *For* $R = \prod_{i=1}^d (a_i, b_i] \in \mathcal{H}$, *show that*

$$\lambda(R) = \lambda(R^0) = \lambda(\overline{R}) = \prod_{i=1}^d (b_i - a_i)\,.$$

Throughout this subsection and elsewhere, $x \in \mathbb{R}^d$ is to be interpreted as a $d \times 1$ matrix, that is, a column vector of length d. The following is the main result of this subsection.

Theorem 5.2.1. *Suppose* $f : \mathbb{R}^d \to \mathbb{R}^d$ *is defined by*

$$f(x) = Ax + b$$

where A *is a* $d \times d$ *non-singular matrix and* $b \in \mathbb{R}^d$, *then*

$$\lambda\left(f(B)\right) = |\det(A)|\lambda(B) \text{ for all } B \in \mathcal{B}(\mathbb{R}^d), \tag{5.2.1}$$

$\det(A)$ *denoting the determinant of* A.

Proceeding towards the proof, we start with defining "vertices" of a rectangle.

Definition 5.2.3. *For* $R = \prod_{i=1}^{d}(a_i, b_i] \in \mathcal{H}$, *the* set of its vertices $v(R)$ *is defined by*

$$v(R) = \prod_{i=1}^{d}\{a_i, b_i\}\,.$$

The proof of Theorem 5.2.1 relies on the following few lemmas, the first of which is stated in more generality than is immediately needed, for a later requirement. Before stating that, the following definition is needed, which is a generalization of Definition 2.6.3.

Definition 5.2.4. *A measure* μ *on* $\big(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\big)$ *is a* Radon measure *if* $\mu(K) < \infty$ *for every compact set* $K \subset \mathbb{R}^d$.

Lemma 5.2.1. *Suppose* $U \subset \mathbb{R}^d$ *is an open set. If* μ_1, μ_2 *are Radon measures on* $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ *such that* $\mu_1(U^c) = \mu_2(U^c) = 0$ *and*

$$\mu_1(R) \le \mu_2(R) \text{ for all } R \in \mathcal{H} \text{ for which } \overline{R} \subset U \text{ and } v(R) \in \mathbb{Q}^d,$$

then $\mu_1(B) \le \mu_2(B)$ *for all* $B \in \mathcal{B}(\mathbb{R}^d)$.

Proof. We shall first show that the given assumption implies

$$\mu_1(R) \le \mu_2(R) \text{ for all } R \in \mathcal{H} \text{ with } v(R) \in \mathbb{Q}^d\,. \tag{5.2.2}$$

Fix $R = \prod_{i=1}^{d}(a_i, b_i] \in \mathcal{H}$ with $v(R) \in \mathbb{Q}^d$. Let $R_{1,1}, \ldots R_{1,2^d}$ be the 2^d rectangles in $\mathcal{H}$ obtained by bisecting each side of R_1, that is, for each $j = 1, \ldots, 2^d$,

$$R_{1,j} = \prod_{i=1}^{d}(\alpha_i, \beta_i],$$

where for every $i = 1, \ldots, d$, either $\alpha_i = a_i$ and $\beta_i = (a_i + b_i)/2$ or $\beta_i = b_i$ and $\alpha_i = (a_i + b_i)/2$. For $n \ge 2$, let $R_{n,1}, \ldots, R_{n,2^{dn}}$ be the rectangles obtained by bisecting the sides of $R_{n-1,1}, \ldots, R_{n-1,2^{(n-1)d}}$. Define

$$B_n = \bigcup_{1 \le j \le 2^{dn}\,:\,\overline{R_{n,j}} \subset U} R_{n,j},\ n \ge 1\,.$$

The given hypothesis, along with the observation $v(R_{nj}) \in \mathbb{Q}^d$, implies

$$\mu_1(B_n) \le \mu_2(B_n),\ n \ge 1\,.$$

Since U is an open set and the diameter of $\overline{R_{n,j}}$ is 2^{-dn} times the diameter of R, it is immediate that

$$B_n \uparrow R \cap U\,.$$

Continuity of measures from below implies

$$\mu_1(R \cap U) \le \mu_2(R \cap U),$$

which in conjunction with the assumption that μ_1, μ_2 are supported on U shows (5.2.2).

Fix $N \in \mathbb{N}$ and define

$$\mathcal{S} = \left\{ \prod_{i=1}^{d} (a_i, b_i] : -N \le a_i \le b_i \le N \text{ and for all } 1 \le i \le d,\ a_i, b_i \in \mathbb{Q} \right\}.$$

Thinking of $\mathcal{S}$ as a collection of subsets of $\Omega_N := (-N, N]^d$, it is immediate that $\mathcal{S}$ is a semi-field on Ω_N. The restrictions of μ_1, μ_2, which are Radon measures, to Ω_N is finite. With the help of standard measure theoretic arguments using the monotone class theorem, for example, those similar to the ones in the proof of Theorem 2.4.1, (5.2.2) can be shown to imply

$$\mu_1(B) \le \mu_2(B) \text{ for all } B \in \mathcal{B}(\mathbb{R}^d),\ B \subset (-N, N]^d.$$

Finally, the observation

$$\lim_{N \to \infty} \mu_i \left(B \cap (-N, N]^d \right) = \mu_i(B),\ i = 1, 2,$$

completes the proof. ☐

An immediate corollary of the above lemma is the following.

Corollary 5.2.1. *If μ_1, μ_2 are Radon measures on $\left(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\right)$ such that*

$$\mu_1(R) = \mu_2(R) \text{ for all } R \in \mathcal{H} \text{ with } v(R) \in \mathbb{Q}^d,$$

then μ_1, μ_2 agree on $\mathcal{B}(\mathbb{R}^d)$.

Proceeding towards the proof of Theorem 5.2.1, the following lemma shows that like in 1-dimension, the Lebesgue measure on $\mathbb{R}^d$ is translation invariant.

Lemma 5.2.2. *For $B \in \mathcal{B}(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$,*

$$\lambda(B + x) = \lambda(B),$$

where $B + x = \{y + x : y \in B\}$.

Proof. Follows from Exc 5.2.3 and Corollary 5.2.1. ☐

The next lemma, taking us closer to the proof of Theorem 5.2.1, tells us that it suffices to check (5.2.1) for one Borel set B with $0 < \lambda(B) < \infty$.

Lemma 5.2.3. *If f and A are as in the statement of Theorem 5.2.1, then there exists $\theta \in \mathbb{R}$ such that*

$$\lambda\left(f(B)\right) = \theta\lambda(B) \text{ for all } B \in \mathcal{B}(\mathbb{R}^d)\,.$$

Proof. Define

$$\theta = \lambda\left(f\left((0,1]^d\right)\right)\,.$$

Setting

$$\mu_1(B) = \lambda(f(B)),\, B \in \mathcal{B}(\mathbb{R}^d),$$

which is a measure because f is a bijection, and

$$\mu_2(B) = \theta\lambda(B),\, B \in \mathcal{B}(\mathbb{R}^d),$$

the claim would follow from Corollary 5.2.1 once it is shown that

$$\lambda\left(f(R)\right) = \theta\lambda(R) \text{ for all } R \in \mathcal{H} \text{ with } v(R) \in \mathbb{Q}^d\,. \tag{5.2.3}$$

For any $B \in \mathcal{B}(\mathbb{R}^d)$ and $y \in \mathbb{R}^d$,

$$\begin{aligned} f(B+y) &= \{A(x+y)+b : x \in B\} \\ &= \{Ax : x \in B\} + (Ay+b) \\ &= f(B) + Ay\,. \end{aligned}$$

Lemma 5.2.2 shows

$$\lambda\left(f(B+y)\right) = \lambda\left(f(B)\right)\,. \tag{5.2.4}$$

Since for $n = 1, 2, \ldots$,

$$\left(0, \frac{1}{n}\right]^d + x = \prod_{i=1}^d \left(x_i, x_i + \frac{1}{n}\right],\; x = (x_1, \ldots, x_d) \in \mathbb{R}^d,$$

(5.2.4) implies

$$\lambda\left(f\left(\left(0, \frac{1}{n}\right]^d\right)\right) = \lambda\left(f\left(\prod_{i=1}^d \left(x_i, x_i + \frac{1}{n}\right]\right)\right)\,. \tag{5.2.5}$$

Recalling that f is a bijection,

$$(0,1]^d = \bigcup_{i_1,\ldots,i_d=1}^{n} \prod_{j=1}^d \left(\frac{i_j}{n}, \frac{i_j}{n} + \frac{1}{n}\right],$$

and the sets on the right hand side above are disjoint, the definition of θ implies

$$\lambda\left(f\left(\left(0, \frac{1}{n}\right]^d\right)\right) = n^{-d}\theta\,. \tag{5.2.6}$$

Additivity of measure implies that for $m_1, \ldots, m_d \in \mathbb{N}$,

$$\begin{aligned}
\lambda\left(f\left(\prod_{i=1}^{d}\left(0, \frac{m_i}{n}\right]\right)\right) &= \sum_{j_1=1}^{m_1} \cdots \sum_{j_d=1}^{m_d} \lambda\left(f\left(\prod_{i=1}^{d}\left(\frac{m_{i_j}-1}{n}, \frac{m_{i_j}}{n}\right]\right)\right) \\
\text{(by (5.2.5) and (5.2.6))} &= \sum_{j_1=1}^{m_1} \cdots \sum_{j_d=1}^{m_d} n^{-d}\theta \\
&= n^{-d}\theta m_1 \ldots m_d \\
&= \theta \prod_{i=1}^{d} \frac{m_i}{n}\,.
\end{aligned}$$

In other words,

$$\lambda\left(f\left(\prod_{i=1}^{d}(0, r_i]\right)\right) = \theta\lambda\left(\prod_{i=1}^{d}(0, r_i]\right),\ r_1, \ldots, r_d \in \mathbb{Q} \cap (0, \infty)\,.$$

Using (5.2.4) and Lemma 5.2.2 once again establishes (5.2.3), from which, the proof follows. ☐

Proof of Theorem 5.2.1. Let $GL(d, \mathbb{R})$ be the collection of all $d \times d$ nonsingular matrices. Lemma 5.2.3 shows that for all $A \in GL(d, \mathbb{R})$, there exists $\phi(A) \in \mathbb{R}$ such that for all $b \in \mathbb{R}^d$,

$$\lambda\left(\{Ax + b : x \in B\}\right) = \phi(A)\lambda(B) \text{ for all } A \in GL(d, \mathbb{R}),\ B \in \mathcal{B}(\mathbb{R}^d)\,. \tag{5.2.7}$$

To complete the proof, all that needs to be shown is $\phi(A) = |\det(A)|$ for all $A \in GL(d, \mathbb{R})$.

An immediate observation is that

$$\phi(A_1 A_2) = \phi(A_1)\phi(A_2),\ A_1, A_2 \in GL(d, \mathbb{R})\,. \tag{5.2.8}$$

Thus the proof would follow once it can be shown that

$$\phi(D) = \det(D) \text{ for a positive definite diagonal matrix } D, \tag{5.2.9}$$

$$\phi(P) = 1 \text{ for an orthogonal matrix } P, \tag{5.2.10}$$

and that any $A \in GL(d, \mathbb{R})$ can be written as

$$A = P_1 D P_2 \tag{5.2.11}$$

where P_1, P_2 are orthogonal matrices and D is a p.d. diagonal matrix.

If D is a $d \times d$ diagonal matrix whose i-th diagonal entry is $c_i > 0$, then it is immediate that

$$\left\{Dx : x \in (0,1]^d\right\} = \prod_{i=1}^{d}(0, c_i]\,.$$

Putting $B = (0, 1]^d$ in (5.2.7) thus shows

$$\phi(D) = c_1 \dots c_d,$$

(5.2.9) follows from which. If P is a $d \times d$ orthogonal matrix, then

$$\{Px : x \in \mathbb{R}^d,\ \|x\| \le 1\} = \{x \in \mathbb{R}^d : \|x\| \le 1\},$$

$\| \cdot \|$ being the usual L^2-norm on $\mathbb{R}^d$. Taking B to be the set on the right hand side above, putting it in (5.2.7) and using the fact $0 < \lambda(B) < \infty$, (5.2.10) follows.

Finally for (5.2.11), the spectral theorem implies

$$AA' = Q_1 W Q_1'$$

for some orthogonal matrix Q_1 and p.d. diagonal matrix W. Let D be the diagonal matrix whose entries are the positive square root of the corresponding entries of W. Define

$$C = Q_1 D Q_1'$$

which is obviously non-singular, and

$$Q_2 = C^{-1} A\,. \tag{5.2.12}$$

Since C is a symmetric matrix, it follows from the above that

$$\begin{aligned} Q_2 Q_2' &= C^{-1} A A' C^{-1} \\ \text{(follows from the definition of } C) &= Q_1 D^{-1} Q_1' (AA') Q_1 D^{-1} Q_1' \\ \text{(choice of } Q_1, W) &= Q_1 D^{-1} Q_1' (Q_1 W Q_1') Q_1 D^{-1} Q_1' \\ \big(Q_1' Q_1 = I\big) &= Q_1 D^{-1} W D^{-1} Q_1' \\ \big(Q_1' Q_1 = I,\ W = D^2\big) &= I, \end{aligned}$$

showing Q_2 is an orthogonal matrix. Rewrite (5.2.12) as

$$\begin{aligned} A &= C Q_2 \\ &= Q_1 D Q_1' Q_2\,. \end{aligned}$$

Setting $P_1 = Q_1$ and $P_2 = Q_1' Q_2$, and observing P_1, P_2 are orthogonal matrices and D is diagonal and p.d., (5.2.11) follows.

Combine (5.2.8) and (5.2.11) to write

$$\begin{aligned} \phi(A) &= \phi(P_1)\phi(D)\phi(P_1) \\ \text{(by (5.2.9) and (5.2.10))} &= \det(D) \\ &= |\det(A)|, \end{aligned}$$

the last line following from (5.2.11) and the facts $\det(UV) = \det(U)\det(V)$ and $\det(P) = \pm 1$ if P is orthogonal. Hence the proof follows. $\square$

Exercise 5.2.4. *Show that* $\lambda(V + x) = 0$ *if* V *is a subspace of* $\mathbb{R}^d$ *with dimension at most* $d - 1$.

5.3 The Jacobian Theorem

For stating the result, a Jacobian matrix has to be first defined. Consider an open set $U \subset \mathbb{R}^d$ and a function $F : U \to \mathbb{R}^d$. Denote by $f_1, \ldots, f_d$ the coordinate functions of F, that is,

$$F(x) = \big(f_1(x), \ldots, f_d(x)\big)\,, x \in \mathbb{R}^d\,.$$

If the first partial derivatives of F exist, that is, $\partial f_i(x)/\partial x_j$ exists for all $x \in U$ and $1 \le i, j \le d$, then its Jacobian matrix at x, denoted by $J(x)$, is a $d \times d$ matrix defined by

$$J(x) = \left(\frac{\partial f_i(x)}{\partial x_j}\right)_{1 \le i,j \le d}, x \in U\,,$$

that is, the (i, j)-th entry of $J(x)$ is $\partial f_i(x)/\partial x_j$. The statement of the theorem is the following, of which, the one-dimensional change of variables formula is a special case. As mentioned in Definition 5.2.1, the integral of a Borel function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ with respect to λ, which is the Lebesgue measure on $\mathbb{R}^d$, is denoted by

$$\int f(x)\, dx,$$

whenever it is defined.

Theorem 5.3.1 (The Jacobian theorem)**.** *For open subsets U and V of $\mathbb{R}^d$, let $T : U \to V$ be a bijection which is continuously differentiable, that is, the first partial derivatives of T exist and are continuous. Assume that its Jacobian matrix $J(x)$ is non-singular for all $x \in U$. Then for any measurable function $f : V \to \overline{\mathbb{R}}$,*

$$\int_U f\big(T(x)\big)\, |\det(J(x))|\; dx = \int_V f(y)\, dy\,,$$

whenever the integral on either side is defined.

Putting $d = 1$ in the above theorem yields the following.

Corollary 5.3.1. *Suppose that U and V are open subsets of $\mathbb{R}$ and $T : U \to V$ is a continuously differentiable bijection whose derivative T' never vanishes. Then*

$$\int_U f\left(T(x)\right) |T'(x)|\, dx = \int_V f(y)\, dy,$$

for a measurable $f : V \to \overline{\mathbb{R}}$ whenever either side above is defined.

Corollary 5.3.1 is slightly weaker than Corollary 4.3.1 because the continuity of the derivative is not assumed in the latter.

The following facts from multivariable analysis are needed for the proof. The first one is the inverse function theorem.

Fact 5.3.1. *Let* $U \subset \mathbb{R}^d$ *be an open set and* $T : U \to \mathbb{R}^d$ *be continuously differentiable. Denoting by* $J(x)$ *the Jacobian matrix of* T *at* $x \in U$*, assume that* $J(x_0)$ *is non-singular for some* $x_0 \in U$*. Then, there exists an open neighbourhood* X *of* x_0 *such that* T *is one-one on* X*, the set* $T(X)$ *is open,* T^{-1} *is continuously differentiable on* $T(X)$ *and the Jacobian matrix of* T^{-1} *at* y *is* $\left(J(T^{-1}y)\right)^{-1}$ *for all* $y \in T(X)$*.*

The following is another fact from multivariable analysis which essentially follows from the one-dimensional mean value theorem.

Fact 5.3.2. *Suppose that* $U \subset \mathbb{R}^d$ *is open,* $R \subset U$ *is a compact rectangle and* $T : U \to \mathbb{R}^d$ *is continuously differentiable such that*

$$|J_{ij}(y) - J_{ij}(x)| \le \alpha \,, x, y \in R \,, 1 \le i, j \le d \,,$$

where $J_{ij}(z)$ *is the* (i, j)*-th entry of the Jacobian matrix* $J(z)$ *of* T *at* z *for all* $z \in U$ *and* $1 \le i, j \le d$*. Then,*

$$\|T(x) - T(y) - J(x)(x - y)\| \le d\alpha\|x - y\| \,, x, y \in R \,,$$

where $\| \cdot \|$ *is the* L^∞ *norm on* $\mathbb{R}^d$ *defined by*

$$\|x\| = \max_{1 \le i \le d} |x_i| \,, x = (x_1, \ldots, x_d) \in \mathbb{R}^d \,, \tag{5.3.1}$$

$x, y, T(x), T(y)$ *are viewed as* $d \times 1$ *vectors and hence* $J(x)(x - y)$ *is also a* $d \times 1$ *vector.*

Proof of Fact 5.3.2. Since $\| \cdot \|$ is the L^∞ norm, it suffices to show that the absolute value of each entry of the $d \times 1$ vector $T(x) - T(y) - J(x)(x - y)$ is at most $d\alpha\|x - y\|$. In other words, it suffices to show that if $f : U \to \mathbb{R}$ is continuously differentiable, and

$$|f_i(y) - f_i(x)| \le \alpha \,, x, y \in R, \; i = 1, \ldots, d \,,$$

where

$$f_i(x) = \frac{\partial f(x)}{\partial x_i} \,, x \in U, \; i = 1, \ldots, d \,,$$

then

$$\left| f(x) - f(y) - \sum_{i=1}^{d} f_i(x)(x_i - y_i) \right| \le d\alpha\|x - y\| \,, x, y \in R \,.$$

Let f be a function satisfying the hypotheses. Let $x^0 = x$, $x^d = y$, and for $1 \le i \le d - 1$,

$$x^i = (y_1, \ldots, y_i, x_{i+1}, \ldots, x_d) \,.$$

Since R is a rectangle, $x^1, \ldots, x^{d-1} \in R$. For a fixed $i = 1, \ldots, d$, x^{i-1} and x^i have all entries identical except the i-th one, which are x_i and y_i

respectively. The one-dimensional mean value theorem implies there exists ξ_i between x_i and y_i such that

$$f\left(x^{i-1}\right) - f\left(x^i\right) = (x_i - y_i) f_i(y_1, \ldots, y_{i-1}, \xi_i, x_{i+1}, \ldots, x_d)\,.$$

Since $\tilde{\xi_i} = (y_1, \ldots, y_{i-1}, \xi_i, x_{i+1}, \ldots, x_d) \in R$ because R is a rectangle, the hypotheses on f imply

$$\left| f_i(\tilde{\xi_i}) - f_i(x) \right| \leq \alpha\,, i = 1, \ldots, d\,.$$

Therefore,

$$\begin{aligned}
&\left| f(x) - f(y) - \sum_{i=1}^{d} f_i(x)(x_i - y_i) \right| \\
&= \left| f\left(x^0\right) - f\left(x^d\right) - \sum_{i=1}^{d} f_i(x)(x_i - y_i) \right| \\
&= \left| \sum_{i=1}^{d} \left[f\left(x^{i-1}\right) - f\left(x^i\right) \right] - \sum_{i=1}^{d} f_i(x)(x_i - y_i) \right| \\
&= \left| \sum_{i=1}^{d} \left(f_i(\tilde{\xi_i}) - f_i(x) \right) (x_i - y_i) \right| \\
&\leq \sum_{i=1}^{d} \left| f_i(\tilde{\xi_i}) - f_i(x) \right| |x_i - y_i| \\
&\leq d\alpha \max_{1 \leq i \leq d} |x_i - y_i| \\
&= d\alpha \|x - y\|\,.
\end{aligned}$$

This completes the proof. □

Proof of Theorem 5.3.1

Without loss of generality, assume f to be non-negative and finite. The proof of Theorem 5.3.1 will be executed by sequentially showing each step below. Step 4. would complete the proof.

Step 1. For any $R \in \mathcal{H}$ with $\overline{R} \subset U$ and $v(R) \in \mathbb{Q}^d$,

$$\lambda\big(T(R)\big) \leq \int_R |\det(J(x))|\, dx\,. \tag{5.3.2}$$

Step 2. For all $A \in \mathcal{B}(\mathbb{R}^d)$,

$$\lambda\big(T(A \cap U)\big) \leq \int_{A \cap U} |\det(J(x))|\, dx\,. \tag{5.3.3}$$

Step 3. For any non-negative measurable function $f : V \to \mathbb{R}$,

$$\int_U f\big(T(x)\big)\, |\det(J(x))|\, dx \ge \int_V f(y)\, dy\,. \tag{5.3.4}$$

Step 4. The inequality in (5.3.4) is an equality.

The proof of Step 1., which is the main step of the proof, is based on the idea that locally T is like a linear transformation.

Proof of Step 1. Fix $R \in \mathcal{H}$ with $\overline{R} \subset U$ and $v(R) \in \mathbb{Q}^d$; $\overline{R}$ is a compact set as R is bounded. Let $\varepsilon > 0$. Since $\det(J(\cdot))$ is a continuous function, it is uniformly continuous on $\overline{R}$. Choose $\delta_1 > 0$ such that

$$\big|\det(J(x)) - \det(J(x'))\big| \le \varepsilon \text{ for all } x, x' \in \overline{R},\ \|x - x'\| \le \delta_1\,, \tag{5.3.5}$$

where $\|\cdot\|$ denotes the L^∞ norm as in (5.3.1) throughout.

Recall that the function $A \to A^{-1}$, from the space of $d \times d$ non-singular matrices to itself, is continuous. Since $J(x)$ is non-singular for all $x \in U$, the map $x \mapsto J(x)^{-1}$ is continuous on U. Thus,

$$f : \overline{R} \times \{z \in \mathbb{R}^d : \|z\| = 1\} \to \mathbb{R}^d\,,$$

defined by

$$f(x, z) = J(x)^{-1} z\,, (x, z) \in \overline{R} \times \{z \in \mathbb{R}^d : \|z\| = 1\}\,,$$

is a continuous function defined on a compact set; elements of $\mathbb{R}^d$ are viewed as $d \times 1$ vectors by convention. Therefore,

$$c = \max\Big\{\|f(x, z)\| : \ (x, z) \in \overline{R} \times \{z \in \mathbb{R}^d : \|z\| = 1\}\Big\} < \infty\,.$$

In other words,

$$\big\|J(x)^{-1} z\big\| \le c\|z\|\,, x \in \overline{R},\ z \in \mathbb{R}^d\,. \tag{5.3.6}$$

Denote by $J_{ij}(x)$ the (i, j)-th entry of $J(x)$ for all $x \in U$ and $1 \le i, j \le d$. Uniform continuity of $J_{ij}(\cdot)$ on $\overline{R}$ ensures the existence of $\delta_2 > 0$ such that

$$\big|J_{ij}(x) - J_{ij}(x')\big| \le \frac{\varepsilon}{cd} \text{ for all } x, x' \in \overline{R},\ \|x - x'\| \le \delta_2\,. \tag{5.3.7}$$

Let $0 < \delta \le \min\{\delta_1, \delta_2\}$ be such that $\delta^{-1}(b_i - a_i)$ is an integer for every i. Choosing such a δ is possible because $b_i - a_i$ is rational; if p_i, q_i are positive integers with $b_i - a_i = p_i/q_i$, letting

$$\delta = \frac{1}{nq_1 \ldots q_d}\,,$$

works for large n, for example.

Consider the square

$$[a_1 + (i_1 - 1)\delta, a_1 + i_1\delta] \times \ldots \times [a_d + (i_d - 1)\delta, a_d + i_d\delta] \ ,$$

where $i_1, \ldots, i_d$ are positive integers with $i_j \le \delta^{-1}(b_j - a_j)$ for $j = 1, \ldots, d$. Denote the collection of all such squares by $\{Q_1, \ldots, Q_k\}$. In other words, $Q_1, \ldots, Q_k$ are compact squares of side-length δ such that

$$\overline{R} = Q_1 \cup \ldots \cup Q_k \, , \tag{5.3.8}$$

and $\lambda(Q_i \cap Q_j) = 0$ for $1 \le i < j \le k$ by Exc 5.2.4. Let x_i be the centre of Q_i (the centre of a square or a rectangle is well defined). Recalling that $\|\cdot\|$ is the L^∞ norm, write

$$Q_i = B_{\delta/2}(x_i)\, , i = 1, \ldots, k\, , \tag{5.3.9}$$

where for $r \ge 0$ and $z = (z_1, \ldots, z_d) \in \mathbb{R}^d$,

$$B_r(z) = \{y \in \mathbb{R}^d : \|y - z\| \le r\} = [z_1 - r, z_1 + r] \times \ldots \times [z_d - r, z_d + r]\, . \tag{5.3.10}$$

The above is precisely the advantage of working with the L^∞ norm.

For $i = 1, \ldots, k$, define

$$\phi_i(z) = J(x_i)(z - x_i) + T(x_i)\, , z \in \mathbb{R}^d\, .$$

Our first claim is that

$$T\left(Q_i\right) \subset \phi_i\left(Q_i^\varepsilon\right)\, , i = 1, \ldots, k\, , \tag{5.3.11}$$

where

$$Q_i^\varepsilon = B_{(1+\varepsilon)\delta/2}(x_i)\, , i = 1, \ldots, k\, .$$

Proceeding towards proving (5.3.11), fix $i \in \{1, \ldots, k\}$, and use Fact 5.3.2 along with (5.3.7) to claim that for all $z \in Q_i$,

$$\|T(z) - T(x_i) - J(x_i)(z - x_i)\| \le \frac{\varepsilon}{c}\|z - x_i\|\, .$$

Since the left hand side above equals $\|T(z) - \phi_i(z)\|$, it follows that

$$\|T(z) - \phi_i(z)\| \le \frac{\varepsilon}{c}\|z - x_i\|\, , z \in Q_i\, . \tag{5.3.12}$$

Therefore, for $z \in Q_i$,

$$\begin{aligned} \left\|\phi_i^{-1} \circ T(z) - z\right\| &= \left\|\phi_i^{-1} \circ T(z) - \phi_i^{-1} \circ \phi_i(z)\right\| \\ &= \left\|J(x_i)^{-1}\left(T(z) - \phi_i(z)\right)\right\| \\ &\le c\left\|T(z) - \phi_i(z)\right\| \\ &\le \varepsilon\|z - x_i\|\, , \end{aligned}$$

(5.3.6) and (5.3.12) implying the inequalities in the penultimate line and the last line, respectively. Thus, for $z \in Q_i$,

$$\left\|\phi_i^{-1} \circ T(z) - x_i\right\| \le \left\|\phi_i^{-1} \circ T(z) - z\right\| + \|z - x_i\| \le (1+\varepsilon)\|z - x_i\|\,.$$

Recall (5.3.9) to argue that

$$\phi_i^{-1} \circ T(z) \in Q_i^\varepsilon\,, z \in Q_i\,,$$

which is equivalent to (5.3.11).

An immediate implication of (5.3.11) is that for fixed $i = 1, \ldots, k$,

$$\begin{aligned}\lambda\left(T(Q_i)\right) &\le \lambda\left(\{J(x_i)z + T(x_i) - J(x_i)x_i : z \in Q_i^\varepsilon\}\right) \\ &= |\det(J(x_i))|\lambda(Q_i^\varepsilon)\,,\end{aligned}$$

the second line following from Theorem 5.2.1. This is the crux of the proof in that it shows how the modulus of the determinant of the Jacobian appears. Further, (5.3.10) shows Q_i^ε is a square of side-length $(1+\varepsilon)\delta$. Therefore,

$$\lambda(Q_i^\varepsilon) = (1+\varepsilon)^d\delta^d = (1+\varepsilon)^d\lambda(Q_i)\,,$$

(5.3.9) implying the second equality. Put everything together to get

$$\lambda\left(T(Q_i)\right) \le |\det(J(x_i))|(1+\varepsilon)^d\lambda(Q_i)\,.$$

Thus,

$$\begin{aligned}\lambda\left(T(R)\right) &\le \sum_{i=1}^k \lambda(T(Q_i)) \\ &\le (1+\varepsilon)^d \sum_{i=1}^k |\det(J(x_i))|\lambda(Q_i) \\ \text{(by (5.3.5) and } \delta \le \delta_1) &\le (1+\varepsilon)^d \sum_{i=1}^k \left(\varepsilon + \min_{z \in Q_i} |\det(J(z))|\right)\lambda(Q_i) \\ (\lambda(Q_i \cap Q_j) = 0,\, i \ne j) &\le (1+\varepsilon)^d\left(\varepsilon\lambda(R) + \int_{Q_1\cup\ldots\cup Q_k} |\det(J(x))|\,dx\right) \\ &= (1+\varepsilon)^d\left(\varepsilon\lambda(R) + \int_R |\det(J(x))|dx\right),\end{aligned}$$

the last line following from (5.3.8) and Exc 5.2.3. Since the above holds for all $\varepsilon > 0$, letting $\varepsilon \downarrow 0$ completes the proof of Step 1. ☐

While Step 1. was mostly based on analysis and linear algebra, the proof of Step 2. is standard in measure theory and follows from Lemma 5.2.1.

Proof of Step 2. Define measures μ and ν on $\mathbb{R}^d$ by

$$\mu(A) = \lambda(T(A \cap U)) \, , A \in \mathcal{B}(\mathbb{R}^d) \, ,$$

and

$$\nu(B) = \int_{B \cap U} |\det(J(x))| \; dx \, , B \in \mathcal{B}(\mathbb{R}^d) \, .$$

The claim (5.3.3) is equivalent to

$$\mu(A) \le \nu(A) \, , A \in \mathcal{B}(\mathbb{R}^d) \, . \tag{5.3.13}$$

In view of Lemma 5.2.1, it suffices to show that the claim holds for any $R \in \mathcal{H}$ with $\overline{R} \subset U$ and $v(R) \in \mathbb{Q}^d$, that is,

$$\mu(R) \le \nu(R) \, . \tag{5.3.14}$$

This is precisely what has been shown in Step 1. □

The proof of Step 3., which is also standard, is based on approximating a non-negative measurable function by simple functions from below.

Proof of Step 3. First let $f : V \to \mathbb{R}$ be a non-negative simple function, that is,

$$f = \sum_{i=1}^k \alpha_i \mathbf{1}_{A_i} \, ,$$

for some $\alpha_1, \ldots, \alpha_k \in [0, \infty]$ and $A_1, \ldots, A_k \in \mathcal{B}(\mathbb{R}^d)$ with $A_i \subset V$ for all i. Then,

$$\begin{aligned} \int_V f(y) \, dy &= \sum_{i=1}^k \alpha_i \lambda(A_i) \\ &= \sum_{i=1}^k \alpha_i \lambda(T(T^{-1} A_i)) \\ &\le \sum_{i=1}^k \alpha_i \int_{T^{-1} A_i} |\det(J(x))| \, dx \\ &= \int_U |\det(J(x))| \sum_{i=1}^k \alpha_i \mathbf{1}_{T^{-1} A_i}(x) \, dx \\ &= \int_U |\det(J(x))| \sum_{i=1}^k \alpha_i \mathbf{1}_{A_i}(T(x)) \, dx \\ &= \int_U |\det(J(x))| \, f(T(x)) \, dx \, , \end{aligned}$$

the inequality in the third line following from Step 2. Thus,

$$\int_V f(y)\,dy \le \int_U |\det(J(x))|\, f(T(x))\,dx\,. \tag{5.3.15}$$

For a measurable function $f : V \to [0,\infty)$, there exist non-negative simple functions f_n such that $f_n \uparrow f$. The desired inequality (5.3.15) holds with f replaced by f_n therein. Letting $n \to \infty$ with the help of MCT, the proof of Step 3. follows. □

Step 4. is a consequence of the inverse function theorem.

Proof of Step 4. Fact 5.3.1 and the assumption that $J(x)$ is non-singular for all $x \in U$ imply that $T^{-1} : V \to U$ is a continuously differentiable bijection whose Jacobian matrix is $J(T^{-1}y)^{-1}$ for all $y \in V$. Using Step 3. with U, V, T replaced by V, U, T^{-1} implies

$$\int_U g(x)\,dx \le \int_V g \circ T^{-1}(y) \left|\det\left(J(T^{-1}y)^{-1}\right)\right| dy\,, \tag{5.3.16}$$

for any measurable $g : U \to [0,\infty)$.

Fix a measurable $f : V \to [0,\infty)$. Define

$$g(x) = f \circ T(x) |\det(J(x))|\,, x \in U\,.$$

Apply (5.3.16) to this g to get

$$\begin{aligned}\int_U f \circ T(x) |\det(J(x))|\,dx &\le \int_V g \circ T^{-1}(y) \left|\det\left(J(T^{-1}y)^{-1}\right)\right| dy \\ &= \int_V f(y) \left|\det\left(J(T^{-1}y)\right)\right| \left|\det\left(J(T^{-1}y)^{-1}\right)\right| dy \\ &= \int_V f(y)\,dy\,.\end{aligned}$$

Compare this with (5.3.4) obtained in Step 3. to get

$$\int_U f \circ T(x) |\det(J(x))|\,dx = \int_V f(y)\,dy\,.$$

This completes the proof of Step 4. and that of Theorem 5.3.1 as well. □

The following special case of Theorem 5.3.1 deserves special mention.

Corollary 5.3.2 (Transformation to polar coordinates)**.** *For a Borel measurable $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$,*

$$\int_{\mathbb{R}^2} f(x,y)\,\lambda(dx,dy) = \int_{(0,2\pi)\times(0,\infty)} f\left(r\cos\theta, r\sin\theta\right) r\,\lambda(dr,\,d\theta), \tag{5.3.17}$$

whenever the integral on either side makes sense.

Proof. Use Theorem 5.3.1 with $U = (0, 2\pi) \times (0, \infty)$, $V = \mathbb{R}^2 \setminus \{(x, 0) : x \geq 0\}$ and $T : U \to V$ defined by

$$T(\theta, r) = (r \cos\theta, r \sin\theta)\,, (\theta, r) \in U\,.$$

The Jacobian matrix of T is

$$J(\theta, r) = \begin{bmatrix} -r \sin\theta & r \cos\theta \\ \cos\theta & \sin\theta \end{bmatrix}, (\theta, r) \in U\,,$$

and hence

$$|\det J(\theta, r)| = r\,.$$

Since T is a continuously differentiable bijection from U to V whose Jacobian matrix is always non-singular, (5.3.17) follows from Theorem 5.3.1 whenever either side of it makes sense. ☐