Universitat de les Illes Balears

FINAL DEGREE REPORT

MEASURE THEORY AND APPLICATIONS

Frank William Hammond Espinosa

Grau de Matemàtiques

Escola Politècnica Superior

Academic year 2023-24

MEASURE THEORY AND APPLICATIONS

Frank William Hammond Espinosa

Final Degree Report

Escola Politècnica Superior

Universitat de les Illes Balears

Academic year 2023-24

Paraules clau del treball: Measure Theory, Integration, Probability Theory

Tutor: Bartomeu Coll Vicens i Maria Jolis Gimenez

Este escrito se lo dedico a mis amigos dentro y fuera del país. En especial, a mi familia y a Marga.

Contents

CHAPTER 1

INTRODUCTION

Metric concepts such as longitudes, areas and volumes come as very natural to us. It seems almost evident that most objects appearing in our daily life can be assigned a number, its *volume*, and the same is true when imagining two-dimensional objects in a plane or one-dimensional objects in a line. They are so intuitive that in the seventeenth century, when calculus appeared, the integral was regarded simply as “the area under a curve”, and little attention was given to what the word “area” meant or how it related to different curves (can every curve be assigned an “area under it”?).

The first attempt at formalising the notion of integral that is studied today was that of Bernhard Riemann. The notion of volume can then be obtained simply as the integral of the constant function 1. His approach is now at the core of most university courses on basic analysis, and it suffices for most real-world applications regarding the calculus of areas and volumes. Mathematically, however, Riemann integration presents some problems, which become apparent when trying to integrate non-continuous functions.

Lebesgue integration and measure theory appeared to try to solve these problems, generalising Riemann integration in almost all cases of interest, and being the foundations for more advanced mathematical theories in analysis. An example of this is functional analysis: L^p or Sobolev spaces are always defined in measure-theoretic terms [1, 2].

This theory, abstract as it is, has many applications to the real world, one of which is image processing: state-of-the-art variational models think of images as the minimisers of some energy functional defined on some function space, always on measure-theoretic terms [3, 4].

Another application of measure theory is found in probability theory: in his groundbreaking work [5], Kolmogorov defined probability theory in terms of the abstract integration and measure theory developed years earlier by Lebesgue, Fréchet et al. His approach has become the standard way of understanding probability theory today.

The goal of this work is to provide a solid background in abstract Measure Theory that allows the study of more complex topics in analysis, as well as developing the theory and language necessary to develop an important theorem in probability theory: the

Kolmogorov Extension Theorem.

The text is structured as follows: the first half of the text is composed entirely by Chapter 2, which is devoted to the development of basic definitions and results in measure theory, as well as the treatment of the abstract Lebesgue integral. Then, tools required for the proof of the Kolmogorov Extension Theorem are detailed in Chapter 3. In Chapter 4, we study a systematic way to talk about measures in higher dimensions, and obtain the Kolmogorov Extension Theorem as a final result. Finally, in Chapter 5, we use all the theory developed so far to obtain some applications to real analysis and probability theory. The formal dependencies between sections are represented in the following figure:

$$
\begin{tikzcd}
1 & & 2.3 \arrow[r] \arrow[ddr] & 3.1 \arrow[r] & 3.2 \arrow[d] & \\
2.1 \arrow[r] & 2.2 \arrow[ur] \arrow[dr] & A \arrow[u] \arrow[d] & & 4.2 \arrow[dr] & \\
 & & 2.4 \arrow[r] \arrow[uur] \arrow[d] & 4.1 \arrow[r] \arrow[ur] & 5.1 \arrow[r] \arrow[d] & 5.2 \\
 & & B.1 \arrow[r] & B.2 & C &
\end{tikzcd}
$$

Figure 1.1: Formal dependencies between the sections of this work. An arrow $X \to Y$ means that a part of section X is needed to study Y.

Most of the content of the text was written following the excellent structure, explanation and exposition of [6]. Any praise on the quality of said book would be an understatement; simply put, it is the perfect material for an undergraduate student who wants to gain some insight into measure theory, real analysis and probability theory. Sometimes, however, the approach followed was original; in those cases, it has been indicated in the text.

It should be noted that the Axiom of Choice is used indiscriminately throughout the text. Two logical equivalences used are Zorn's Lemma (in the proof of the Radon-Nikodým Theorem) and the statement that every (infinite) cartesian product of nonempty set is nonempty (in Section 4.2).

CHAPTER 2

BASIC MEASURE THEORY

2.1 Notation

Let $\mathbb{R}$ be the set of real numbers. We will define the **extended real line** in the usual way; that is, as the set $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$, where $+\infty$ and $-\infty$ are formal symbols for which we extend the order and arithmetic of $\mathbb{R}$ in the following manner:

$$\forall a \in \mathbb{R}\colon\ -\infty < a < +\infty$$
$$+\infty + (+\infty) = +\infty, \quad -\infty + (-\infty) = -\infty, \quad 0 \cdot (+\infty) = (+\infty) \cdot 0 = 0 \cdot (-\infty) = (-\infty) \cdot 0 = 0$$
$$\forall a \in \mathbb{R}\colon\ -\infty + a = a + (-\infty) = -\infty, \quad +\infty + a = a + (+\infty) = +\infty, \quad \frac{a}{+\infty} = \frac{a}{-\infty} = 0$$
$$\forall b \in \overline{\mathbb{R}} \setminus \{0\}:\ (+\infty) \cdot b = b \cdot (+\infty) = \begin{cases} +\infty, & b > 0 \\ -\infty, & b < 0 \end{cases}, (-\infty) \cdot b = b \cdot (-\infty) = \begin{cases} -\infty, & b > 0 \\ +\infty, & b < 0 \end{cases}$$

We say that $a < \infty$ whenever $a \neq -\infty$ and $a \neq +\infty$. If there is no confusion with the previous notation, the symbol $+\infty$ is often also denoted as just ∞. The expressions $+\infty + (-\infty), -\infty + (+\infty), \frac{+\infty}{+\infty}, \frac{-\infty}{+\infty}, \frac{+\infty}{-\infty}$ and $\frac{-\infty}{-\infty}$ are not defined. We say that an expression or arithmetic operation in $\overline{\mathbb{R}}$ is **well-defined** (or, simply, **defined**), if it is not one of the above undefined expressions. We also extend the notion of **interval** to $\overline{\mathbb{R}}$ in an intuitive way: given $a, b \in \overline{\mathbb{R}}$, define $[a, b] = \{x \in \overline{\mathbb{R}}\colon a \le x \le b\}$, $(a, b) = \{x \in \overline{\mathbb{R}}\colon a < x < b\}$ and other kinds of intervals similarly. The class of all intervals of the form (a, b), $[-\infty, b)$ or $(a, +\infty]$ with $a, b \in \overline{\mathbb{R}}$, $a \le b$, forms the basis of a topology on $\overline{\mathbb{R}}$, which we will call the **standard topology** on $\overline{\mathbb{R}}$. Intervals of the form $(a, b]$ or $[a, b]$ with $a, b \in \overline{\mathbb{R}}$, $a \le b$, will be of special interest, and they will be called **right-semiclosed** intervals.

Let $A, A_1, \ldots, A_n, \ldots$ be subsets of some nonempty set Ω. We say that the sets A_n form an **increasing** sequence of sets whenever $A_1 \subseteq \cdots \subseteq A_n \subseteq \ldots$ If $A = \bigcup_n A_n$, we denote it by $A_n \uparrow A$. We say that the sets A_n form a **decreasing** sequence of sets whenever $A_1 \supseteq \cdots \supseteq A_n \supseteq \ldots$ If $A = \bigcap_n A_n$, we denote it by $A_n \downarrow A$. We denote their **upper** and **lower limits** as $\limsup_n A_n = \bigcap_n \bigcup_{k \ge n} A_k$ and $\liminf_n A_n = \bigcup_n \bigcap_{k \ge n} A_k$, respectively.

Similarly, if $f_1, f_2, \ldots$ form an increasing (decreasing) sequence of (extended) real-valued functions with limit f, we may write $f_n \uparrow f$ ($f_n \downarrow f$). The same convention is used for monotone sequences of (extended) real numbers.

Sometimes, if a sequence of sets, functions or numbers has more than one index, (i.e., $\{A_{nm}\}_{n,m\in\mathbb{N}}$) and is monotone with respect to one, the index is specified as a subindex of the arrow (i.e., $A_{nm}\uparrow_n B_m$, where $B_m=\bigcup_n A_{nm}$).

If f and g are functions from some set Ω to $\overline{\mathbb{R}}$, statements such as $f\le g$ or $f=g$ are always interpreted pointwise, that is, $f(\omega)\le g(\omega)$ or $f(\omega)=g(\omega)$ for all $\omega\in\Omega$. The same is true for expressions like the limit of a sequence of functions or its supremum.

We say that a function f is **positive** (resp., **negative**) if $f>0$ ($f<0$), and **nonnegative** (**nonpositive**) if $f\ge 0$ ($f\le 0$). If $f\colon\Omega\to\overline{\mathbb{R}}$, its **positive** and **negative parts** are defined by $f^+(\omega)=\max(f(\omega),0)$, $f^-(\omega)=-\min(f(\omega),0)=\max(-f(\omega),0)$.

The notation for $\{f\ge 0\}$ (and similar expressions) is interpreted likewise: $\{f\ge 0\}=\{\omega\in\Omega\,|\,f(\omega)\ge 0\}$. Also, if $B\subseteq\overline{\mathbb{R}}$, we write $\{f\in B\}=\{\omega\in\Omega\,|\,f(\omega)\in B\}=f^{-1}(B)$.

Finally, if $A\subseteq\Omega$, the **indicator function** of A is defined as $I_A(\omega)=1$ if $\omega\in A$ and $I_A(\omega)=0$ otherwise.

2.2 Introductory results and definitions

In this section, we introduce the basic concepts regarding measure theory: fields, σ-fields and measures, among others. As we will see, there is a strong algebraic component in measure theory. For instance, it is not possible to assign an *intuitive*[1] measure every subset of $\mathbb{R}$ (see 2.18 in [7]), so it is necessary to find suitable structures on which to develop the theory. This will be the role played by the concepts introduced here:

Definition 2.2.1. *Let Ω be an arbitrary set and let $\mathscr{F}$ be a class of subsets of Ω. We say that $\mathscr{F}$ is*

- *A **field** or **algebra** if $\emptyset\in\mathscr{F}$ and it is closed under finite unions and complementation, that is, if we consider any two $A,B\in\mathscr{F}$, then $A^c\in\mathscr{F}$ and $A\cup B\in\mathscr{F}$.*
- *A **σ-field** or **σ-algebra** if $\emptyset\in\mathscr{F}$ and it is closed under countable unions and complementation, that is, if we consider any sequence of sets $A_1,A_2,\dots\in\mathscr{F}$, then $A_1^c\in\mathscr{F}$ and $\bigcup_n A_n\in\mathscr{F}$.*
- ***Monotone** if it is closed under monotone sequences, that is, if $A_n\in\mathscr{F}$ and $A_n\downarrow A$ or $A_n\uparrow A$, then $A\in\mathscr{F}$.*

From this definition in follows easily that a field is closed under finite intersection, a σ-field is closed under countable intersection and that a class of sets is a σ-field if, and only if, it is both monotone and a field.

It is easy to check that the arbitrary intersection of fields, σ-fields or monotone classes is, respectively, a field, σ-field or monotone class. This allows us to ensure the existence of the respective structure containing a given class $\mathscr{S}$ of subsets of Ω. We shall call them the **minimal** field, σ-field or monotone class over $\mathscr{S}$ or say that they are **generated by** $\mathscr{S}$. They will be written, respectively, as $\mathscr{F}(\mathscr{S}),\sigma(\mathscr{S})$ and $\mathscr{C}(\mathscr{S})$. We say that $\mathscr{S}$

[1] More concretely, if using the axiom of choice, it is not possible to define a measure on the power set $\mathscr{P}(\mathbb{R})$ that assigns its length to each interval.

generates $\mathscr{F}(\mathscr{S}), \sigma(\mathscr{S})$ and $\mathscr{C}(\mathscr{S})$. We will often used reasonings involving this concept. One that is almost self-evident - but very useful and common - is the following:

Remark 2.2.2. *Let $\mathscr{S}$ be a class of subsets of Ω. If $\mathscr{A}$ is another class of subsets of Ω that has a given structure and contains $\mathscr{S}$, then the respective structure generated by $\mathscr{S}$ is also contained in $\mathscr{A}$; say, $\mathscr{A}$ is a σ-field. Then, $\mathscr{S} \subseteq \mathscr{A}$ implies $\sigma(\mathscr{S}) \subseteq \mathscr{A}$.*

Let X be a topological space. The class of **Borel sets** of X, denoted by $\mathscr{B}(X)$, is the smallest σ-field containing all open sets of X. The Borel sets of $\mathbb{R}$ and $\overline{\mathbb{R}}$ are of special interest, and having a small enough class of generators will be very convenient.

Proposition 2.2.3. *The class of Borel sets of $\mathbb{R}$ is generated by the class of all intervals of a specified form. Namely, every family of intervals of one of the following forms:*

(i) $(-\infty, b)$, $b \in \mathbb{R}$ *(ii)* $(a, +\infty)$, $a \in \mathbb{R}$ *(iii)* (a, b), $a \le b \in \mathbb{R}$ *(iv)* $[a, b]$, $a \le b \in \mathbb{R}$
(v) $(a, b]$, $a \le b \in \mathbb{R}$ *(vi)* $[a, b)$, $a \le b \in \mathbb{R}$ *(vii)* $(-\infty, b]$, $b \in \mathbb{R}$ *(viii)* $[a, +\infty)$, $a \in \mathbb{R}$

Proof. We will first prove the result for open, bounded intervals. Let $\mathscr{F}$ denote the σ-field generated by all open intervals, and let $\mathscr{T}$ be the standard topology on $\mathbb{R}$ (that is, the class of all open sets). It is known that the cartesian product of countable sets is countable. Thus, every subset of $\mathbb{Q} \times \mathbb{Q}$ is either finite or countable.
Let $U \in \mathscr{T}$. Now note that, by the density of the rationals and the fact that U is open,

$$U = \bigcup_{a,b} (a, b),$$

where a and b range over the pairs of rational numbers such that $a < b$ and $(a, b) \subseteq U$ (which there is, at most, countably many of). Hence, $\mathscr{T} \subseteq \mathscr{F}$. Since $\mathscr{F}$ is a σ-field containing $\mathscr{T}$, by Remark 2.2.2, $\sigma(\mathscr{T}) \subseteq \mathscr{F}$. But, by definition, $\mathscr{B}(\mathbb{R}) = \sigma(\mathscr{T})$. Reciprocally, since $\mathscr{B}(\mathbb{R})$ contains all open intervals, the smallest σ-field containing all open intervals, $\mathscr{F}$, satisfies $\mathscr{F} \subseteq \mathscr{B}(\mathbb{R})$.
This completes the proof for open, bounded intervals. To see it for other kinds of intervals, simply note that any interval can be expressed by finite or countable unions and intersections of any other given kind of intervals or their complements. For instance, $(a, b] = \bigcap_n (a, b + 1/n)$ or $[a, b) = [a, +\infty) \cap ([b, +\infty))^c$. □

Proposition 2.2.4. *The class of Borel sets of $\overline{\mathbb{R}}$ is generated by the class of all intervals of one of these forms:*

(i) $[-\infty, b]$, $b \in \overline{\mathbb{R}}$ *(ii)* $[-\infty, b)$, $a \in \overline{\mathbb{R}}$ *(iii)* $(a, +\infty]$, $a \in \overline{\mathbb{R}}$ *(iv)* $[a, +\infty]$, $a \in \overline{\mathbb{R}}$

Proof. Note that the singletons $\{-\infty\}$ and $\{+\infty\}$ are closed in $\overline{\mathbb{R}}$, since their complements $(-\infty, +\infty]$ and $[-\infty, +\infty)$ are open. It is also not hard to see that every open subset of $\mathbb{R}$ is open if regarded as a subset of $\overline{\mathbb{R}}$.
Define $\mathscr{M} = \left\{B \in \mathscr{B}\left(\overline{\mathbb{R}}\right) | B \setminus \{-\infty, +\infty\} \in \mathscr{B}(\mathbb{R})\right\}$. It is now easy to see that $\mathscr{M}$ is a σ-field containing all open sets of $\overline{\mathbb{R}}$, hence $\mathscr{B}\left(\overline{\mathbb{R}}\right) \subseteq \mathscr{M}$. Since $\mathscr{M} \subseteq \mathscr{B}\left(\overline{\mathbb{R}}\right)$ by definition, it follows that B is a Borel set of $\overline{\mathbb{R}}$ if, and only if, $B \setminus \{-\infty, +\infty\}$ is a Borel set of $\mathbb{R}$; hence, $C \in \mathscr{B}(\mathbb{R})$ if, and only if, $C, C \cup \{-\infty\}, C \cup \{+\infty\}$ and $C \cup \{-\infty, +\infty\} \in \mathscr{B}\left(\overline{\mathbb{R}}\right)$.
Let $\mathscr{I}$ be any of the classes of sets (i)-(iv), and denote $\mathscr{I}' = \{B \setminus \{-\infty, +\infty\} | B \in \mathscr{I}\}$. It is clear that $\sigma(\mathscr{I}) \subseteq \mathscr{B}\left(\overline{\mathbb{R}}\right)$, and by Proposition 2.2.3, $\sigma(\mathscr{I}') = \mathscr{B}(\mathbb{R})$. Finally, define

$\mathcal{M}' = \{C \in \mathcal{B}(\mathbb{R}) \,|\, C, C \cup \{-\infty\}, C \cup \{+\infty\} \text{ and } C \cup \{-\infty, +\infty\} \in \sigma(\mathcal{I})\}$. It is clear, by the form of $\mathcal{I}$, that $\mathcal{M}'$ is a σ-field containing $\mathcal{I}'$; hence, $\mathcal{M}' = \mathcal{B}(\mathbb{R})$. This implies, however, that $\mathcal{B}(\overline{\mathbb{R}}) \subseteq \sigma(\mathcal{I})$. □

Definition 2.2.5. *Let Ω be a nonempty set and $\mathcal{S}$ a class of subsets of Ω. A **set function** on $\mathcal{S}$ (or, simply, a set function) is a mapping $\lambda\colon \mathcal{S} \to \overline{\mathbb{R}}$. We say that a set function λ is **finitely additive**, or simply **additive**, if the values $+\infty$ and $-\infty$ do not both belong to the image of λ, there exists some $A \in \mathcal{S}$ such that $\lambda(A)$ is finite[2] and*

$$\lambda\left(\bigcup_n A_n\right) = \sum_n \lambda(A_n) \tag{2.1}$$

*for every finite family of disjoint sets $A_1, A_2, \ldots \in \mathcal{S}$ such that $\bigcup_n A_n \in \mathcal{S}$ (this is always the case if $\mathcal{S}$ is a field). If condition (2.1) instead holds for every countable family of subsets whose union belongs to $\mathcal{S}$ (this will always be the case if $\mathcal{S}$ is a σ-field), we say that λ is **countably additive** or **σ-additive**[3].*

*If $\mathcal{S}$ is a σ-field, then a nonnegative, countably additive set function μ is called a **measure** on $\mathcal{S}$. A measure satisfying $\mu(\Omega) = 1$ is called a **probability measure** or, simply, a **probability**.*

Definition 2.2.6. *A **measurable space** is a pair $(\Omega, \mathcal{F})$, where Ω is a nonempty set and $\mathcal{F}$ is a σ-field of subsets of Ω. A **measure space** is a tuple $(\Omega, \mathcal{F}, \mu)$, where $(\Omega, \mathcal{F})$ is a measurable space and μ is a measure on $\mathcal{F}$. A **probability space** is a measure space $(\Omega, \mathcal{F}, p)$ where p is a probability measure on $\mathcal{F}$.*

Proposition 2.2.7. *Let λ be a finitely additive set function on the field $\mathcal{F}_0$. Then,*

(i) $\lambda(\emptyset) = 0$

(ii) $\lambda(A \cup B) + \lambda(A \cap B) = \lambda(A) + \lambda(B)$ *for all* $A, B \in \mathcal{F}_0$.

(iii) If $A, B \in \mathcal{F}_0$ and $A \subseteq B$, then $\lambda(B) = \lambda(A) + \lambda(B \setminus A)$. In particular, $\lambda(B) \geq \lambda(A)$ if $\lambda(B \setminus A) \geq 0$ and $\lambda(B \setminus A) = \lambda(B) - \lambda(A)$ if $\lambda(A) < \infty$.

(iv) If λ is nonnegative,

$$\lambda\left(\bigcup_{k=1}^n A_k\right) \leq \sum_{k=1}^n \lambda(A_k) \quad \textit{for all } A_1, \ldots, A_n \in \mathcal{F}_0$$

(v) If λ is a measure,

$$\lambda\left(\bigcup_n A_n\right) \leq \sum_n \lambda(A_n)$$

for all $A_1, A_2, \cdots \in \mathcal{F}$ such that $\bigcup_n A_n \in \mathcal{F}$.

[2]This condition is not imposed in some literature, allowing the set functions $\lambda_1(A) = +\infty$ and $\lambda_2(A) = -\infty$ to be counted as additive, since they satisfy all other requirements, but they are degenerate cases and will be excluded in this text. The main reason why we consider them degenerate cases is that $\lambda_i(\emptyset) \neq 0$. On the contrary, if there exists some A such that $\lambda(A) < \infty$, then $\lambda(\emptyset) = 0$ (see Proposition 2.2.7.(i)).

[3]Note that a necessary condition for a σ-additive function to be well-defined is that for every sequence of sets $A_1, A_2, \ldots$ such that $\bigcup_n A_n \in \mathcal{F}$, $\forall n : \lambda(A_n) < \infty$ and $\lambda(\bigcup_n A_n) < \infty$, the series of real numbers $\sum_n \lambda(A_n)$ is absolutely convergent, because $\bigcup_n A_n$ is invariant under permutations of indices, while the series $\sum_n \lambda(A_n)$ is only invariant under permutations of indices if it is absolutely convergent.

Proof. 1. Take any set $A \in \mathscr{F}$ such that $\lambda(A) < \infty$. Then,

$$\lambda(A) = \lambda(A \cup \emptyset) = \lambda(A) + \lambda(\emptyset),$$

whence $\lambda(\emptyset) = 0$.

2. Note that $\lambda(A \cup B) = \lambda(A \setminus B) + \lambda(B \setminus A) + \lambda(A \cap B)$. Therefore,

$$\lambda(A \cup B) + \lambda(A \cap B) = (\lambda(A \setminus B) + \lambda(A \cap B)) + (\lambda(B \setminus A) + \lambda(A \cap B)) = \lambda(A) + \lambda(B).$$

3. Immediate by additivity.

4. Write $B_n = A_n \setminus (A_1 \cup \cdots \cup A_{n-1}) \in \mathscr{F}$. Since $B_n \subseteq A_n$, by 2.2.7.(iii), $\lambda(B_n) \leq \lambda(A_n)$. Note that the sets B_n are disjoint and their union is $\bigcup_n A_n$. Thus,

$$\lambda\left(\bigcup_n A_n\right) = \sum_n \lambda(B_n) \leq \sum_n \lambda(A_n).$$

5. The proof given for 2.2.7.(iv) still holds word for word (now the union is infinite, but the notation used is the same). □

Definition 2.2.8. *A set function λ defined on a class $\mathscr{S}$ of subsets of Ω is said to be* ***finite*** *if $\lambda(A) < \infty$ for every $A \in \mathscr{S}$. If $\mathscr{F}_0$ is a field and λ is finitely additive, it suffices that $\lambda(\Omega)$ be finite, for $\lambda(\Omega) = \lambda(A) + \lambda(A^c)$; and if $\lambda(A)$ is infinite, so is $\lambda(\Omega)$.*
A nonnegative, finitely additive set function λ on a field $\mathscr{F}_0$ is said to be ***σ-finite*** *whenever Ω can be written as $\bigcup_n A_n$, where $A_n \in \mathscr{F}_0$*[4] *and $\lambda(A_n) < \infty$ for every n.*

Remark 2.2.9. *Let λ be a finitely additive set function on a field $\mathscr{F}_0$. Then, consider $A, B \in \mathscr{F}_0$ such that $A \subseteq B$. If $\lambda(A) = \pm\infty$, 2.2.7.(iii) implies that $\lambda(B) = \lambda(A) + \lambda(B \setminus A) = \pm\infty + \lambda(B \setminus A) = \pm\infty$. As a consequence, if $\lambda(B) < \infty$, then $\lambda(A) < \infty$.*
Another interesting thing to note is that, by Proposition 2.2.7.(iii), every finite measure is bounded. We will see later on that this is the case too for countably additive set functions.

One of the most common processes in analysis is taking limits. The following definition and the subsequent two propositions, which will be of great usefulness during the rest of the text, relate this process to the language we have been developing.

Definition 2.2.10. *A set function λ defined on some class of subsets $\mathscr{S}$ is said to be* ***continuous from below*** *at a given $A \in \mathscr{S}$ whenever $\lim_n \lambda(A_n) = \lambda(A)$ for every increasing sequence of sets $A_n \uparrow A$, with $A_n \in \mathscr{S}$ for all n. Is is said to be* ***continuous from above*** *at a given $A \in \mathscr{S}$ whenever $\lim_n \lambda(A_n) = \lambda(A)$ for every decreasing sequence of sets $A_n \downarrow A$, with $A_n \in \mathscr{S}$ for all n.*

Proposition 2.2.11. *Let λ be a σ-additive set function on a field $\mathscr{F}_0$. Then,*

[4]The condition that $A_n \in \mathscr{F}_0$ is important. The following scenario will be quite common in the rest of the text: we have a σ-field $\mathscr{F}$, and a field $\mathscr{F}_0$ such that $\mathscr{F} = \sigma(\mathscr{F}_0)$, and we have information on $\mathscr{F}_0$ that we want to extend to the whole σ-field $\mathscr{F}$. It is a requirement for some theorems (see the Carathéodory Extension Theorem) that a measure is σ-finite specifically over $\mathscr{F}_0$, and not over the whole σ-field $\mathscr{F}$.

(i) *λ is continuous from below at every $A \in \mathscr{F}_0$; that is, if $A_n \uparrow A$ and $A \in \mathscr{F}_0$, then* $\lim_n \lambda(A_n) = \lambda(A)$.

(ii) *λ is continuous from above at every $A \in \mathscr{F}_0$ with $\lambda(A) < \infty$ if we only consider decreasing sequences $A_1, A_2, \ldots \in \mathscr{F}_0$ such that $\lambda(A_1) < \infty$. More concretely : if $A_n \downarrow A$, $A \in \mathscr{F}_0$ and $\lambda(A_1) < \infty$, then* $\lim_n \lambda(A_n) = \lambda(A)$.

Proof. 1. Define $B_1 = A_1$, $B_n = A_n \setminus (A_1 \cup \ldots \cup A_{n-1}) = A_n \setminus A_{n-1} \in \mathscr{F}_0$, so that the sets B_n are disjoint and $A_n = B_1 \cup \ldots \cup B_n$. Therefore, by additivity, $\lambda(A_n) = \sum_{k=1}^n \lambda(B_k)$. Finally, since $A = \bigcup_n B_n$,

$$\lim_n \lambda(A_n) = \sum_{k=1}^{+\infty} \lambda(B_k) = \lambda(A)$$

2. Define $C_n = A_1 \setminus A_n$. Since $A_n \subseteq A_1$ and $A \subseteq A_1$, by Remark 2.2.9, $\lambda(A_n) < \infty$ and $\lambda(A) < \infty$. Thus, by Proposition 2.2.7.(iii), $\lambda(C_n) = \lambda(A_n) - \lambda(A_1)$ and $\lambda(A_1 \setminus A) = \lambda(A_1) - \lambda(A)$. The desired result now follows from 2.2.11.(i) taking into consideration that $C_n \uparrow (A_1 \setminus A)$. □

We can state a result which is in some way reciprocal to the previous one:

Proposition 2.2.12. *Let λ be a finitely additive function defined on a field $\mathscr{F}_0$. Then,*

(i) *If λ is continuous from below at every $A \in \mathscr{F}_0$, then it is σ-additive on $\mathscr{F}_0$.*

(ii) *If λ is continuous from above at the empty set, then it is σ-additive on $\mathscr{F}_0$.*

Proof. 1. Let $A_1, A_2, \ldots \in \mathscr{F}_0$ be a sequence of disjoint sets such that $\bigcup_n A_n \in \mathscr{F}_0$. Define $A = \bigcup_n A_n$ and $B_n = A_1 \cup \ldots \cup A_n \in \mathscr{F}_0$. Then, $B_n \uparrow \bigcup_n A_n$ and, by additivity, $\lambda(B_n) = \sum_{k=1}^n \lambda(A_k)$. Since λ is continuous from below at $\bigcup_n A_n$, we have

$$\lambda\left(\bigcup_n A_n\right) = \lim_n \lambda(B_n) = \lim_n \sum_{k=1}^n \lambda(A_k) = \sum_n \lambda(A_n).$$

2. We will show that λ is continuous from below at every $A \in \mathscr{F}_0$: let $A_1, A_2, \ldots \in \mathscr{F}_0$ be a sequence of sets increasing to $A \in \mathscr{F}_0$. Define $B_n = A \setminus A_n$. It is clear that $B_n \downarrow \emptyset$. Additionally, $\lambda(B_n) + \lambda(A_n) = \lambda(A)$. Since $\lambda(B_n) \to 0$, it must be $\lambda(A_n) \to \lambda(A)$. By 2.2.12.(i), λ is σ-additive. □

2.3 Extension of measures

The goal of this section is to extend, under certain technical hypotheses that will appear later on, a nonnegative, σ-additive set function μ over a field $\mathscr{F}_0$ into a measure over a σ-field that contains $\mathscr{F}_0$. Although this may seem artificial to the reader at the moment, it is very common in measure theory to find the need of extending a measure in this way. To this avail, we shall follow the following scheme:

- First, we restrict ourselves to the case where μ is finite, which, up to a rescaling, is equivalent to it being a probability measure. Then, we extend μ to the class of countable unions of sets of $\mathscr{F}_0$, $\mathscr{C}$, by taking limits. This collection is closer to being a σ-field, but it need not contain complements of sets in it.

- Secondly, we extend the function obtained to all subsets of Ω, via approximating them by sets we can measure (sets in $\mathscr{C}$).

- This extension will turn out not to be a measure in all of $\mathscr{P}(\Omega)$, but we can find a subset where it *is* a measure, and said subset will turn out to be a σ-field containing $\mathscr{F}_0$.

- Finally, we are able to drop the finiteness restriction over μ and cover a more general case, by using the construction above. Moreover, the construction made will allow us to conclude that said extension is, in fact, unique.

Having a general overview of the ideas followed, let us now dive into the details.

Lemma 2.3.1. *Let $\mathscr{F}_0$ be a field of subsets of a given set Ω, and let p be a nonnegative, countably additive set function on $\mathscr{F}_0$ such that $p(\Omega) = 1$. Let $A_1, A_2, \ldots$ be family of sets that belong to $\mathscr{F}_0$ and increase to a limit A. Take $A'_1, A'_2, \ldots$ and A' similarly (note that A and A' need not belong to $\mathscr{F}_0$). If $A \subseteq A'$, then*

$$\lim_n p(A_n) \leq \lim_n p(A'_n).$$

Proof. Firstly, note that both limits exist since $\{p(A_n)\}_n$ and $\{p(A'_n)\}_n$ are both increasing sequences of real numbers, and bounded by 1, following Proposition 2.2.7.(iii).
Take $m \in \mathbb{Z}^+$. Then, $A_m \cap A'_n \uparrow_n A_m \cap A' = A_m \in \mathscr{F}$. Thus, by Proposition 2.2.11.(i),

$$\lim_n p(A_m \cap A'_n) = p(A_m).$$

But $p(A_m \cap A'_n) \leq p(A'_n)$ (Proposition 2.2.7.(iii)), whence

$$p(A_m) \leq \lim_n p(A'_n).$$

The result follows easily from taking limits when $m \to \infty$ in the above expression. □

Now we can extend p to a larger class of sets: the class of countable unions of sets of $\mathscr{F}_0$.

Lemma 2.3.2. *Let $\mathscr{C}$ be the class of all countable unions of sets in $\mathscr{F}_0$. Define μ on $\mathscr{C}$ as follows: if $A \in \mathscr{C}$, there exists a sequence of sets $A_n \in \mathscr{F}_0$ increasing to A. Now set $\mu(A) = \lim_n p(A_n)$. This limit exists since the sequence is increasing and bounded by* 1 *and μ is well-defined by Lemma 2.3.1. Also, clearly $\mu \equiv p$ on $\mathscr{F}_0$. Then:*

(i) $\emptyset, \Omega \in \mathscr{C}$*,* $\mu(\emptyset) = 0$*,* $\mu(\Omega) = 1$ *and* $0 \leq \mu(A) \leq 1$ *for all* $A \in \mathscr{C}$

(ii) If $G_1, G_2 \in \mathscr{C}$*, then* $G_1 \cup G_2, G_1 \cap G_2 \in \mathscr{C}$ *and* $\mu(G_1 \cup G_2) + \mu(G_1 \cap G_2) = \mu(G_1) + \mu(G_2)$*.*

(iii) *If $G_1, G_2 \in \mathscr{C}$ and $G_1 \subseteq G_2$, then $\mu(G_1) \le \mu(G_2)$.*

(iv) *If $G_n \in \mathscr{C}$, and $G_n \uparrow G$, then $G \in \mathscr{C}$ and $\mu(G_n) \to \mu(G)$.*

Proof. 1. Follows easily from the fact that $\mu \equiv p$ in $\mathscr{F}_0$, taking into account that p is a probability measure.

2. Let $A_n^1 \uparrow G_1$, $A_n^1 \in \mathscr{F}_0$, and $A_n^2 \uparrow G_2$, $A_n^2 \in \mathscr{F}_0$. Then, $A_n^1 \cup A_n^2 \uparrow G_1 \cup G_2$ and $A_n^1 \cap A_n^2 \uparrow G_1 \cap G_2$. We have

$$\forall n \in \mathbb{Z}^+ : p(A_n^1 \cup A_n^2) + p(A_n^1 \cap A_n^2) = p(A_n^1) + p(A_n^2).$$

The proof is completed by taking limits in the above expression.

3. This follows easily from Lemma 2.3.1.

4. For each $m \in \mathbb{Z}^+$, take $A_{m,n} \uparrow_n G_m$, $A_{n,m} \in \mathscr{F}_0$. To see that $G \in \mathscr{C}$, simply take any bijection $\rho \colon \mathbb{Z}^+ \to \mathbb{Z}^+ \times \mathbb{Z}^+$ and note that $G = \bigcup_{k \in \mathbb{Z}^+} A_{\rho(k)}$. Now define, for each $m \in \mathbb{Z}^+$, $D_m = A_{1,m} \cup A_{2,m} \cup \cdots \cup A_{m,m} \in \mathscr{F}_0$. Note that, by definition, $\forall n : A_{n,m} \subseteq D_m$, and $D_m \subseteq G_1 \cup G_2 \cup \cdots \cup G_m = G_m$. Therefore, the following inclusion holds for every n:

$$A_{n,m} \subseteq D_m \subseteq G_m, \tag{2.2}$$

and by 2.3.2.(iii), we have $\mu(A_{n,m}) \le \mu(D_m) \le \mu(G_m)$, which can be written as

$$p(A_{n,m}) \le p(D_m) \le \mu(G_m) \tag{2.3}$$

Let $m \to +\infty$ in (2.2) to see that $G_n \subseteq \bigcup_{m \in \mathbb{Z}^+} D_m \subseteq G$ and now let $n \to +\infty$ to see that $D_m \uparrow G$, and thus $\mu(G) = \lim_m p(D_m)$. Finally, let $m \to +\infty$ in (2.3) to see that $\mu(G_n) \le \mu(G) \le \lim_m \mu(G_m)$ and now let $n \to +\infty$ to complete the proof. □

We now are going to extend μ to the class of all subsets of Ω. This construction only depends on properties $(i)-(iv)$ in Lemma 2.3.2 and not on the initial definition of μ.

Lemma 2.3.3. *Let $\mathscr{C}$ be a class of subsets of a given set Ω, μ a nonnegative real-valued set function on $\mathscr{C}$ such that $\mathscr{C}$ and μ satisfy the four conditions $(i)-(iv)$ of Lemma 2.3.2. Define, for each $A \subseteq \Omega$,*

$$\mu^*(A) = \inf\{\mu(G) : G \in \mathscr{C}, A \subseteq G\}.$$

Then,

(i) $\mu^* \equiv \mu$ *on $\mathscr{C}$, and* $0 \le \mu^*(A) \le 1$ *for all $A \subseteq \Omega$.*

(ii) $\mu^*(A \cup B) + \mu^*(A \cap B) \le \mu^*(A) + \mu^*(B)$.

(iii) $\mu^*(A) + \mu^*(A^c) \ge 1$.

(iv) *If $A \subseteq B$, then $\mu^*(A) \le \mu^*(B)$.*

(v) *If $A_n \uparrow A$, then $\mu^*(A_n) \to \mu^*(A)$.*

Proof. 1. Take any set $A \in \mathscr{C}$. Then, for all $G \in \mathscr{C}, A \subseteq G$, we have $\mu(A) \leq \mu(G)$ by 2.3.2.(iii). This lower bound is achieved by A itself, and thus $\mu^*(A) = \mu(A)$. Bounds 0 and 1 follow easily from the definition of infimum.

2. For any $\varepsilon > 0$, take $G_1, G_2 \in \mathscr{C}$ such that $A \subseteq G_1, B \subseteq G_2$ and $\mu(G_1) \leq \mu^*(A) + \varepsilon/2, \mu(G_2) \leq \mu^*(B) + \varepsilon/2$. Therefore,

$$\mu^*(A)+\mu^*(B)+\varepsilon \geq \mu(G_1)+\mu(G_2) = \mu(G_1 \cup G_2)+\mu(G_1 \cap G_2) \geq \mu^*(A \cup B)+\mu^*(A \cap B).$$

Since $\varepsilon > 0$ is arbitrary, the result holds.

3. Immediate consequence of 2.3.3.(ii): $\mu^*(A) + \mu^*(A^c) \geq \mu^*(\Omega) + \mu^*(\emptyset) = 1$, where $\mu^*(\emptyset) = 0$ by 2.3.3.(i).

4. Immediate by definition.

5. First, note that $\lim_n \mu^*(A_n)$ exists (we have an increasing sequence of real numbers bounded by 1 following 2.3.3.(i) and 2.3.3.(iv)) and is bounded above by $\mu^*(A)$, because $\forall n : \mu^*(A_n) \leq \mu^*(A)$ by 2.3.3.(iv). Let $\varepsilon > 0$, and define $\varepsilon_n = \varepsilon/2^n$ (this choice is made so that $\sum_n \varepsilon_n = \varepsilon$). Consider $G_n \in \mathscr{C}$ such that $A_n \subseteq G_n$ and $\mu(G_n) \leq \mu^*(A_n) + \varepsilon_n$. Now, $A = \bigcup_n A_n \subseteq \bigcup_n G_n$, whence

$$\mu^*(A) \leq \mu^*\left(\bigcup_n G_n\right) = \mu\left(\bigcup_n G_n\right) = \lim_n \mu\left(\bigcup_{k=1}^n G_n\right),$$

where in the last step we used Lemma 2.3.2.(iv). If we can show that $\mu\left(\bigcup_{k=1}^n G_k\right) \leq \mu^*(A_n) + \varepsilon$, the proof will be done. With this in mind, it suffices to prove that $\mu\left(\bigcup_{k=1}^n G_k\right) \leq \mu^*(A_n) + \sum_{k=1}^n \varepsilon_k$ for all $n \in \mathbb{Z}^+$, since $\sum_{k=1}^n \varepsilon_k < \sum_k \varepsilon_k = \varepsilon$. The case $n = 1$ is true by construction. Now apply Lemma 2.3.2.(ii) to $A = \bigcup_{k=1}^n G_k$ and $B = G_{k+1}$:

$$\mu\left(\bigcup_{k=1}^{n+1} G_k\right) = \mu\left(\left(\bigcup_{k=1}^{n} G_k\right) \cup G_{n+1}\right) = \mu\left(\bigcup_{k=1}^{n} G_k\right) + \mu(G_{n+1}) - \mu\left(\bigcup_{k=1}^{n} G_k \cap G_{n+1}\right).$$

Note that $A_n = A_n \cap A_{n+1} \subseteq G_n \cap G_{n+1} \subseteq \bigcup_{k=1}^n G_k \cap G_{n+1}$, and that implies $\mu^*(A_n) \leq \mu\left(\bigcup_{k=1}^n G_k \cap G_{n+1}\right)$. Moreover, $\mu(G_{n+1}) \leq \mu^*(A_{n+1}) + \varepsilon_{n+1}$. Using both inequalities and the induction hypothesis,

$$\mu\left(\bigcup_{k=1}^{n+1} G_k\right) \leq \mu^*(A_n) + \sum_{k=1}^{n} \varepsilon_k + \mu^*(A_{n+1}) + \varepsilon_{n+1} - \mu(A_n) = \mu^*(A_{n+1}) + \sum_{k=1}^{n+1} \varepsilon_k,$$

thus completing the proof. □

As discussed in the beginning of the section, this extension will, in general, not be a measure over all of $\mathscr{P}(\Omega)$. However, we can find a suitable σ-field for it to be a measure on, following a somewhat intuitive idea:
Were μ^* to be a measure over a σ-field $\mathscr{H}$, it would be additive, and for each $A \in \mathscr{H}$,

$$\mu^*(A) + \mu^*(A^c) = \mu^*(\emptyset) + \mu^*(\Omega) = 1. \tag{2.4}$$

Therefore, we can attempt to define $\mathscr{H}$ as the class of subsets of Ω that satisfy (2.4).

Theorem 2.3.4. *Under the hypotheses of Lemma 2.3.3, let*

$$\mathscr{H} = \{A \subseteq \Omega : \mu^*(A) + \mu^*(A^c) = 1\}$$

(Following Lemma 2.3.3.(iii), the defining property of $\mathscr{H}$ is equivalent to the weaker condition $\mu^*(A)+\mu^*(A^c) \le 1$). Then, $\mathscr{H}$ is a σ-field containing $\mathscr{C}$ and μ^* is a probability measure on $\mathscr{H}$.

Proof. First, note that $\mathscr{C} \subseteq \mathscr{H}$ by Lemma 2.3.2.(ii). We will show that $\mathscr{H}$ is a field. Clearly, it is closed under complementation. Let $H_1, H_2 \in \mathscr{H}$. Then, by Lemma 2.3.3.(ii),

$$\begin{aligned} \mu^*(H_1 \cup H_2) + \mu^*(H_1 \cap H_2) \le \mu^*(H_1) + \mu^*(H_2) \\ \mu^*(H_1^c \cup H_2^c) + \mu^*(H_1^c \cap H_2^c) \le \mu^*(H_1^c) + \mu^*(H_2^c) \end{aligned} \tag{2.5}$$

Adding both inequalities and taking into account that $H_1, H_2 \in \mathscr{H}$, we get that

$$\mu^*(H_1 \cup H_2) + \mu^*((H_1 \cup H_2)^c) + \mu^*(H_1 \cap H_2) + \mu^*((H_1 \cap H_2)^c) \le 2.$$

Define $U = \mu^*(H_1 \cup H_2) + \mu^*((H_1 \cup H_2)^c)$ and $I = \mu^*(H_1 \cap H_2) + \mu^*((H_1 \cap H_2)^c)$ Following, 2.3.3.(iii) $U, I \ge 1$, which means that $2 \le U + I \le 2$, whence $U = I = 1$. It follows that $H_1 \cup H_2 \in \mathscr{H}$, $H_1 \cap H_2 \in \mathscr{H}$. Furthermore, equalities in (2.5) hold, for if inequalities were strict, so would be the right inequality in $2 \le U + I \le 2$. In the case when H_1 and H_2 are disjoint, the first equality degenerates into $\mu^*(H_1 \cup H_2) = \mu^*(H_1) + \mu^*(H_2)$). This shows both that $\mathscr{H}$ is a field and that μ^* is additive on it.
Now consider the countable case. Let $A_n \in \mathscr{H}$, $A_n \uparrow A$ (it is enough to consider this case since $\mathscr{H}$ is a field). Note that $A^c \subseteq A_n^c$, so

$$\mu^*(A_n) + \mu^*(A^c) \le \mu^*(A_n) + \mu^*(A_n^c) = 1.$$

Taking limits, and following Lemma 2.3.3.(v), $\mu^*(A) + \mu^*(A^c) \le 1$. Thus, $A \in \mathscr{H}$. Moreover, μ^* is countably additive in $\mathscr{H}$ by Lemma 2.3.3.(v) and Proposition 2.2.12.(i). □

We can now state our first extension theorem.

Theorem 2.3.5. *A finite measure μ on a field $\mathscr{F}_0$ extends to a measure on $\sigma(\mathscr{F}_0)$.*

Proof. Scale the measure to a probability measure by considering $p = \mu/\mu(\Omega)$. Apply the construction made through Lemma 2.3.2 to Theorem 2.3.4 to obtain a σ-field $\mathscr{H}$ containing $\mathscr{F}_0$ and extend p to a measure on $\mathscr{H}$, $\overline{p}$. Then, $\mu(\Omega) \cdot \overline{p}$ is an extension of μ to $\mathscr{H}$. Since $\mathscr{H}$ is a σ-field containing $\mathscr{F}_0$, we have $\sigma(\mathscr{F}_0) \subseteq \mathscr{H}$, and we can restrict the extension to $\sigma(\mathscr{F}_0)$. □

We have developed all the tools we need to prove the existence part of our more general extension theorem. However, a few extra results can be juiced out of the construction, and some additional tools are required to prove uniqueness. This is why we are going to introduce the concept of *completeness*:

Definition 2.3.6. *Let $(\Omega, \mathscr{F}, \mu)$ be a measure space. A set $A \subseteq \Omega$ is said to be **null** whenever $A \in \mathscr{F}$ and $\mu(A) = 0$. The measure space is said to be **complete** if every subset of a null set is measurable (and, therefore, null too).*
*Consider a measure space $\left(\Omega, \mathscr{F}, \mu\right)$, and let $\mathscr{N}$ be the class of all subsets of null sets in $(\Omega, \mathscr{F}, \mu)$ (i.e. sets A such that $A \subseteq B$ for some null set $B \in \mathscr{F}$). Define $\mathscr{F}_\mu$ as the class of all sets of the form $A \cup N$, where $A \in \mathscr{F}$ and $N \in \mathscr{N}$, and extend μ to $\mathscr{F}_\mu$ as $\mu(A \cup N) = \mu(A)$. The measure space[5] $(\Omega, \mathscr{F}_\mu, \mu)$ is said to be the **completion** of $(\Omega, \mathscr{F}, \mu)$.*

It is not complicated to show that if a measure is complete, then it is its own completion. Along these lines, an interesting thing to note about the completion of a measure space is that it is minimal in the following sense:

Remark 2.3.7. *Let $(\Omega, \mathscr{F}_1, \mu_1)$ and $(\Omega, \mathscr{F}_2, \mu_2)$ be measure spaces over the same set Ω. We say that $(\Omega, \mathscr{F}_2, \mu_2)$ extends $(\Omega, \mathscr{F}_1, \mu_1)$ whenever $\mathscr{F}_1 \subseteq \mathscr{F}_2$ and $\mu_2|_{\mathscr{F}_1} \equiv \mu_1$. Then, every complete extension of a measure space also extends its completion.*

With this in mind, we can show a curious relationship between the σ-field given by Theorem 2.3.4 and $\sigma(\mathscr{F}_0)$:

Proposition 2.3.8. *In Theorem 2.3.4, $(\Omega, \mathscr{H}, \mu^*)$ is the completion of $(\Omega, \sigma(\mathscr{F}_0), \mu^*)$.*

Proof. Let $(\Omega, \mathscr{T}, \mu^*)$ denote the completion of $(\Omega, \sigma(\mathscr{F}_0), \mu^*)$. We shall show that $\mathscr{H}$ extends $\mathscr{T}$ by showing it is complete (see Remark 2.3.7), since it clearly extends $\sigma(\mathscr{F}_0)$. Let $A \in \mathscr{H}$ be a null set, and let $B \subseteq A$. By monotonicity, $\mu^*(B) = 0$, and therefore $\mu^*(B) + \mu^*(B^c) = \mu^*(B^c) \le 1$. Therefore, $B \in \mathscr{H}$.
Reciprocally, let $A \in \mathscr{H}$. Following the definition of μ^*, for every $n \in \mathbb{Z}^+$ there exists some $G_n \in \mathscr{C}$ such that $A \subseteq G_n$, $\mu^*(G_n) - 1/n \le \mu^*(A)$. Similarly, there exists some $G'_n \in \mathscr{C}$ such that $A^c \subseteq G'_n$ and $\mu^*(G'_n) - 1/n \le \mu^*(A^c)$. Let $H_n = G'^c_n$. Then, since $A \in \mathscr{H}$ and $\mathscr{C} \subseteq \mathscr{H}$, $1 - \mu^*(H_n) - 1/n \le 1 - \mu^*(A)$. Thus,

$$\mu^*(G_n) - 1/n \le \mu^*(A) \le \mu^*(H_n) + 1/n.$$

Let $H = \bigcup_n H_n \in \sigma(\mathscr{F}_0)$ and $G = \bigcap_n G_n \in \sigma(\mathscr{F}_0)$. It is clear that $H \subseteq A \subseteq G$. Thus, we can write $A = H \cup (H \setminus A)$, with $H \setminus A \subseteq G \setminus H \in \sigma(\mathscr{F}_0)$, and since $G \setminus H \subseteq G_n \setminus H_n$,

$$\mu^*(G \setminus H) \le \mu^*(G_n \setminus H_n) = \mu^*(G_n) - \mu^*(H_n) \le 2/n.$$

It follows that $\mu^*(G \setminus H) = 0$, and thus $A \in \mathscr{T}$. □

We need one last tool to be able to prove our extension theorem:

[5] Indeed, $\mathscr{F}_\mu$ is a σ-field: it is clearly closed under countable union, and closed under complementation: if $A \in \mathscr{F}_0$ and $N \subseteq M$, $M \in \mathscr{F}$, $\mu(M) = 0$, then $(A \cup N)^c = (M^c \cap A^c) \cup (M \setminus (A \cup N))$, where $A^c \cap M^c \in \mathscr{F}$ and $M \setminus (A \cup N) \subseteq M$. Also, μ is well-defined on $\mathscr{F}_\mu$: if $A_1 \cup N_1 = A_2 \cup N_2$ with $N_i \subseteq M_i$, $\mu(M_i) = 0$, then $A_1 \subseteq A_1 \cup N_1 = A_2 \cup N_2 \subseteq A_2 \cup M_2$, whence

$$\mu(A_1) \le \mu(A_2 \cup M_2) \le \mu(A_2) + \mu(M_2) = \mu(A_2).$$

Similarly, $\mu(A_2) \le \mu(A_1)$. Checking that μ satisfies the axioms of a measure on $\mathscr{F}_\mu$ is simple.

Theorem 2.3.9 (Monotone Class Theorem). *Let $\mathscr{F}_0$ be a field of subsets of Ω. Then, $\mathscr{C}(\mathscr{F}_0) = \sigma(\mathscr{F}_0)$. In particular, if $\mathscr{C}$ is a monotone class of subsets of Ω and $\mathscr{F}_0 \subseteq \mathscr{C}$, then $\sigma(\mathscr{F}_0) \subseteq \mathscr{C}$.*

Proof. Let $\mathscr{M} = \mathscr{C}(\mathscr{F}_0)$ and $\mathscr{F} = \sigma(\mathscr{F}_0)$. Firstly, note that since $\mathscr{F}$ is a monotone class, then $\mathscr{M} \subseteq \mathscr{F}$. Let $A \in \mathscr{F}$, and define $\mathscr{M}_A = \{B \in \mathscr{M} : A \cap B, A \cap B^c \textit{ and } A^c \cap B \in \mathscr{M}\}$. Then, $\mathscr{M}_A$ is a monotone class itself. The rest of the proof is structured as follows:

1. For all $A \in \mathscr{F}_0$, $\mathscr{M}_A = \mathscr{M}$: we have $\mathscr{F}_0 \subseteq \mathscr{M}_A$, and by minimality of $\mathscr{M}$, $\mathscr{M} \subseteq \mathscr{M}_A$. By definition, $\mathscr{M}_A \subseteq \mathscr{M}$. Hence, $\mathscr{M}_A = \mathscr{M}$

2. If B is a set in $\mathscr{M}$ (not necessarily in $\mathscr{F}_0$), then $\mathscr{M}_B = \mathscr{M}$: if $A \in \mathscr{F}_0$, since $\mathscr{M}_A = \mathscr{M}$, we have $A \cap B$, $A \cap B^c$ and $A^c \cap B \in \mathscr{M}$. It follows that $\mathscr{F}_0 \subseteq \mathscr{M}_B$, and this implies that $\mathscr{M} = \mathscr{C}(\mathscr{F}_0) \subseteq \mathscr{M}_B$ (see Remark 2.2.2). $\mathscr{M}_B \subseteq \mathscr{M}$ by definition of $\mathscr{M}_B$.

3. $\mathscr{M}$ is a field because $\emptyset \in \mathscr{F}_0 \subseteq \mathscr{M}$ and it is closed under finite union and complementation because $\mathscr{M} = \mathscr{M}_B$ for every $B \in \mathscr{M}$. Since $\mathscr{M}$ is both a field and a monotone class, it is a σ-field.

Since $\mathscr{M}$ is a σ-field containing $\mathscr{F}_0$, we have $\mathscr{F} = \sigma(\mathscr{F}_0) \subseteq \mathscr{M}$ (again, see Remark 2.2.2). □

We are now ready to prove the main theorem of this section.

Theorem 2.3.10 (Carathéodory Extension Theorem). *Let $\mathscr{F}_0$ be a field of subsets of Ω and μ a nonnegative, countably additive set function over $\mathscr{F}_0$. Assume that μ is **σ-finite** over $\mathscr{F}_0$, that is, that Ω can be written as $\bigcup_n A_n$, where $A_n \in \mathscr{F}_0$ and $\mu(A_n) < +\infty$. Then, μ has a unique extension to the minimal σ-field over $\mathscr{F}_0$.*

Proof. Let $\mathscr{F} = \sigma(\mathscr{F}_0)$. Suppose, without loss of generality, that the sets A_n are disjoint (otherwise, take $A'_n = A_n \setminus \bigcup_{k=1}^{n-1} A_k$; then $\mu(A'_n) \leq \mu(A_n) < +\infty$). Define μ_n in $\mathscr{F}_0$ as $\mu_n(B) = \mu(A_n \cap B)$. Then, μ_n is a finite, nonnegative, countably additive set function in $\mathscr{F}_0$ and can therefore be extended to a measure on $\mathscr{F}$, which we will denote by μ_n^*. Now extend μ to $\mathscr{F}$ by setting $\mu^* = \sum_n \mu_n^*$. Then, μ^* is a measure because the order of summation of any double series of positive terms can be switched (see Corollary A.1.2). To see that μ^* is unique, suppose that λ is also a measure that extends μ to $\mathscr{F}$. Define $\lambda_n(B) = \lambda(B \cap A_n)$. Let $\mathscr{C}_n$ be the class of subsets of Ω where μ_n^* and λ_n coincide. Clearly, $\mathscr{F}_0 \subseteq \mathscr{C}_n$. Moreover, $\mathscr{C}_n$ is a monotone class: if $A_m \uparrow A$ or $A_m \downarrow A$, since both μ_n^* and λ_n are finite, then $\mu_n^*(A) = \lim_m \mu_n^*(A_m) = \lim_m \lambda_n(A_m) = \lambda_n(A)$, whence $A \in \mathscr{C}_n$. By the Monotone Class Theorem, $\lambda_n \equiv \mu_n^*$. Therefore,

$$\lambda = \sum_n \lambda_n = \sum_n \mu_n^* = \mu^*,$$

finishing the proof. □

2.4 Integration

We are now in position of defining the flagship of measure theory: the Lebesgue integral[6]. However, in order to be able to integrate functions, we first need to establish which functions are to be integrated. The concept which will be used reminds that of continuity[7].

Definition 2.4.1. *Let $(\Omega_1, \mathscr{F}_1)$ and $(\Omega_2, \mathscr{F}_2)$ be measurable spaces. A function $f\colon \Omega_1 \to \Omega_2$ is said to be $\mathscr{F}_1, \mathscr{F}_2$-**measurable** whenever $f^{-1}(B) \in \mathscr{F}_1$ for all $B \in \mathscr{F}_2$. Sometimes, one or both σ-fields can be omitted (and, for instance, simply say that f is* measurable*) if it is clear on which σ-fields we are working at.*
When Ω_2 is a topological space, f is said to be Borel measurable (on $(\Omega_1, \mathscr{F}_1)$) if $\mathscr{F}_2 = \mathscr{B}(\Omega_2)$ and f is $\mathscr{F}_1, \mathscr{F}_2$-measurable.

Borel measurable functions taking (extended) real values will be our candidate functions to be integrated. Henceforth, if a function with codomain $\mathbb{R}$ or $\overline{\mathbb{R}}$ is said to be measurable, it will be understood that it is Borel measurable. Before defining the integral, we need to enunciate a series of simple lemmas needed to prove deeper results:

Lemma 2.4.2. *The composition of measurable functions is measurable.*

Proof. Let $(\Omega_1, \mathscr{F}_1)$, $(\Omega_2, \mathscr{F}_2)$ and $(\Omega_3, \mathscr{F}_3)$ be measure spaces and $f\colon \Omega_1 \to \Omega_2$, $g\colon \Omega_2 \to \Omega_3$ measurable functions. Then, for every $M \in \Omega_3$,

$$\left(f \circ g\right)^{-1}(M) = g^{-1}\left(f^{-1}(M)\right).$$

This set is an element of $\mathscr{F}_1$, since $f^{-1}(M) \in \mathscr{F}_2$ because f is measurable and g is measurable. □

Consider the class of all measurable spaces. This class forms a category whose morphisms are measurable functions. This lemma is stating the composability of these morphisms. Identity morphisms are defined as identity functions and associativity follows from associativity of function composition. Said category is sometimes denoted as **Meas**.

Lemma 2.4.3. *Let $(\Omega_1, \mathscr{F}_1)$ and $(\Omega_2, \mathscr{F}_2)$ be measure spaces such that $\mathscr{F}_2 = \sigma(\mathscr{S})$ for some class $\mathscr{S}$ of subsets of Ω_2. Let $h : \Omega_1 \to \Omega_2$. Then, h is measurable if, and only if, $h^{-1}(C) \in \mathscr{F}_1$ for all $C \in \mathscr{S}$.*

Proof. One implication is trivial. To see the other, define

$$\mathscr{M} = \left\{A \in \mathscr{F}_2 \colon h^{-1}(A) \in \mathscr{F}_1\right\}.$$

By hypothesis, $\mathscr{S} \subseteq \mathscr{M}$. Since $\mathscr{F}_1$ is a σ-field and h^{-1} preserves arbitrary unions, intersections and complements, $\mathscr{M}$ is a σ-field. Thus, $\mathscr{F}_2 = \sigma(\mathscr{C}) \subseteq \mathscr{M}$, and therefore h is measurable. □

[6]This is actually the historical motivation behind measure theory: see [8], the original paper by Henri Lebesgue introducing his theory of integration. Measure theory was developed afterwards, to formalise, generalise and justify this new type of integration.

[7]This will be important later on, when we study the relation between topology and measure theory.

This result will show to be very useful, particularly in combination with Proposition 2.2.4 if one wants to show that a given function $h\colon \Omega \to \overline{\mathbb{R}}$ is Borel measurable, it suffices to show, for instance, that $\{h \geq c\}$ (or $\{h < c\}$) is measurable for every $c \in \overline{\mathbb{R}}$.

Lemma 2.4.4. *Let $\Omega, \mathscr{F}$ be a measurable space and f, g Borel measurable functions. Then,* $\max(f,g)$ *defined by* $\max(f,g)(\omega) = \max(f(\omega), g(\omega))$ *is Borel measurable. Similarly,* $\min(f,g)$ *is Borel measurable.*

Proof. By Lemma 2.4.3, it suffices to see that

$$\left\{\omega\colon \max(f,g)(\omega) \leq c\right\} = \left\{\omega\colon f(\omega) \leq c\right\} \cap \left\{\omega\colon g(\omega) \leq c\right\}$$

is measurable for all $c \in \overline{\mathbb{R}}$. Similarly,

$$\left\{\omega\colon \min(f,g)(\omega) \geq c\right\} = \left\{\omega\colon f(\omega) \geq c\right\} \cap \left\{\omega\colon g(\omega) \geq c\right\}$$

is so for all $c \in \overline{\mathbb{R}}$. □

Two last results that are very easy to prove are that for any set $A \subseteq \Omega$, A is measurable if, and only if, its indicator function is Borel measurable; and that every constant function between measure spaces is measurable. Following this last result and Lemma 2.4.4, if h is a Borel measurable function, so are h^+ and h^-.

Proposition 2.4.5. *The pointwise limit of Borel measurable functions is Borel measurable.*

Proof. Let $(\Omega, \mathscr{F})$ be a measurable space and $\{h_n\}_{n\in\mathbb{Z}^+}$ be a sequence of Borel measurable functions $h_n\colon \Omega \to \overline{\mathbb{R}}$ converging pointwise to a limit h. By Lemma 2.4.3, it suffices to show that $\{h > c\} = \{\omega \in \Omega \,|\, h(\omega) > c\}$ is measurable for all c. Now, a simple analytical argument shows that if a_n is a converging sequence of extended real numbers, then for any $c \in \overline{\mathbb{R}}$, $\lim_n a_n > c$ if, and only if, there exist some $r, n_0 \in \mathbb{Z}^+$ such that $a_n > c + \frac{1}{r}$ for any $n \geq n_0$. Therefore,

$$\{h > c\} = \left\{\lim_n h_n > c\right\} = \bigcup_{r\in\mathbb{Z}^+} \bigcup_{n_0\in\mathbb{Z}^+} \bigcap_{n\geq n_0} \left\{h_n > c + \frac{1}{r}\right\}$$

Since $\left\{h_n > c + \frac{1}{r}\right\}$ is measurable for every n and every c, the proof is completed. □

A special kind of Borel measurable functions will be of interest. Concretely, those whose range is a finite set (that is, they take finitely many values). These functions are interesting because their integral can be defined intuitively, and, as we will see, they are closed under arithmetic operations and are able to "generate" all other measurable functions via limits. This will allow us to both define the integral of Borel measurable functions and, later on, find conditions to exchange the integral and limit signs.

Definition 2.4.6 (Simple functions). *Let $(\Omega, \mathscr{F})$ be a measurable space and $h\colon \Omega \to \overline{\mathbb{R}}$ a Borel measurable function. Then, h is said to be a* ***simple function*** *whenever it takes finitely many values. Equivalently, h is simple whenever we can find finitely many measurable sets $A_i \subseteq \Omega$ and values $x_i \in \overline{\mathbb{R}}$ such that*

$$h = \sum_{i=1}^{n} x_i I_{A_i},$$

where I_{A_i} is the indicator function of A_i and the sum is defined, in the sense that the expression $\infty - \infty$ never occurs.

Remark 2.4.7. *If we impose the sets A_i to form a partition of the set Ω, then every simple function h with range $h(\Omega) = \{x_1, \dots, x_n\}$ can be written uniquely as*

$$h = \sum_{i=1}^{n} x_i I_{h^{-1}(x_i)}.$$

*We will call this expression the **standard form** of h*[8].

As we will see, Borel measurable functions are closed under arithmetic operations. However, in order to prove it we need the following lemma.

Lemma 2.4.8. *Any pointwise operation of finitely many simple functions, if defined, is simple. More precisely, if op is an arbitrary mapping $op\colon D \to \overline{\mathbb{R}}$, where $D \subseteq \overline{\mathbb{R}}^n$, and $s_1, \dots, s_n$ are simple functions such that the function $h(\omega) = op(s_1(\omega), \dots, s_n(\omega))$ is well-defined, then h is simple.*[9]

Proof. Write, for any $k \le n$, s_k in standard form as $s_k = \sum_{m=1}^{N_k} x_{km} I_{A_{km}}$ (that is, the sets A_{km} form a partition of Ω). Set $\mathscr{I} = \{1, \dots, N_1\} \times \cdots \times \{1, \dots, N_n\}$, and for every $(m_1, \dots, m_n) \in \mathscr{I}$, define $C_{(m_1,\dots,m_n)} = \bigcap_{k=1}^{n} A_{km_k}$, and $\mathscr{I}' = \left\{\phi \in \mathscr{I} : C_\phi \neq \emptyset\right\}$. Note that, by definition, the family of sets C_ϕ such that $\phi \in \mathscr{I}'$ forms a partition of Ω.
Now, for any $\phi = (m_1, \dots, m_n) \in \mathscr{I}'$, define $x_\phi = op(x_{1m_1}, \dots, x_{nm_n})$. Such value x_ϕ exists because h is defined and for any value $\omega \in C_\phi$ (at least one exists because $\phi \in \mathscr{I}'$, hence C_ϕ is nonempty), then $h(\omega) = op(s_1(\omega), \dots, s_n(\omega)) = op(x_{1m_1}, \dots, x_{nm_n})$.
Using the notation established earlier, we can write

$$h = \sum_{\phi \in \mathscr{I}'} x_\phi I_{C_\phi}.$$

which is clearly a simple function, because $\mathscr{I}' \subseteq \mathscr{I}$ is finite. □

Corollary 2.4.9. *The sum, product and quotient of simple functions, if defined, is a simple function.*

Now we will introduce a theorem that is central for the construction of the Lebesgue integral. It allows us to approximate every measurable function via a sequence of simple functions satisfying useful properties.

Theorem 2.4.10. *(i) Every nonnegative measurable function is the pointwise limit of an increasing sequence of nonnegative, real-valued simple functions. Moreover, if the function is bounded, convergence is uniform.*

(ii) Every measurable function h is the pointwise limit of a sequence of finite-valued simple functions s_n which satisfy $|s_n| \le |h|$ for all $n \in \mathbb{Z}^+$.

[8]This notation is not common in the literature, but it will show to be useful in our text.
[9]Not much attention is given to this in [6]. Both the statement and the proof of this result are original.

Proof. 1. We want to approximate, pointwise, the function h by simple functions s_n. We have almost no information about the domain Ω except from the fact that h is measurable, so we need to work with the codomain $\mathbb{R}$. The idea is to group the *values* the function h takes into finitely many intervals for every fixed value of n, and then recover the subsets of Ω where h takes these values.

For that purpose, equally split the interval $[0,n)$ into $N(n)$ many consecutive intervals $V_k^n = \left[\frac{k-1}{N(n)}, \frac{k}{N(n)}\right)$, $k = 1,\dots,nN(n)$. In this interval, approximate h by its left endpoint $\frac{k-1}{N(n)}$. That is, set

$$s_n(\omega) = \frac{k-1}{N(n)} \quad \text{whenever } h(\omega) \in V_k^n,$$

and $s_n(\omega) = n$ whenever $h(\omega) \geq n$. This way, s_n is nonnegative, finite-valued and $|h - s_n| \leq \frac{1}{N(n)}$ if $h(\omega)$ is in the interval $[0,n]$.

With this construction, each s_n is simple: we can write

$$s_n = nI_{\{h \geq n\}} + \sum_{k=1}^{nN(n)} \frac{k-1}{N(n)} I_{\{h \in V_k^n\}},$$

which is clearly a simple function.

If we want s_n to converge to h, it is sufficient that $N(n)$ tend to $+\infty$ when $n \to +\infty$. This also ensures that convergence is uniform when h is bounded. The main remaining desired condition is that the sequence of functions is increasing. If we impose that it is, we obtain a series of inequalities depending on where $h(\omega)$ falls regarding intervals V_k^n and $V_{k'}^{n+1}$:

- If $h(\omega) \geq n+1$, then $s_n(\omega) = n \leq n+1 = s_{n+1}(\omega)$. This inequality holds for any $N(n)$.
- If $n \leq h(\omega) < n+1$, then $s_n(\omega) = n$ and there exists some k' such that $h(\omega) \in V_{k'}^{n+1}$. This implies that $n \leq h(\omega) < \frac{k'}{N(n+1)}$, from which one can deduce that $k' > nN(n+1)$, hence $k' - 1 \geq nN(n+1)$. Thus,

  $$s_{n+1}(\omega) = \frac{k'-1}{N(n+1)} \geq \frac{nN(n+1)}{N(n+1)} = n = s_n(\omega).$$

  This inequality holds for any value $N(n)$ too.
- If $0 \leq h(\omega) < n$, then $h(\omega) \in V_k^n \cap V_{k'}^{n+1}$ for some k, k'. This implies that

  $$\frac{k-1}{N(n)} \leq h(\omega) < \frac{k'}{N(n+1)},$$

  from which one deduces the inequality $k' > \frac{N(n+1)}{N(n)}(k-1)$. If we impose $\frac{N(n+1)}{N(n)}$ to be a positive integer, it follows that $k' - 1 \geq \frac{N(n+1)}{N(n)}(k-1)$. Therefore,

  $$s_{n+1}(\omega) = \frac{k'-1}{N(n+1)} \geq \frac{k-1}{N(n)} = s_n(\omega).$$

So far, the only conditions imposed to $N(n)$ have been that $\frac{N(n+1)}{N(n)}$ is a positive integer and that $\lim_n N(n) = +\infty$. There are infinitely many ways to do this, but the easiest one is by setting $N(n) = 2^n$.[10]

2. Decompose $h = h^+ - h^-$. Approximate h^+ and h^- by increasing sequences of nonnegative, finite-valued, simple functions s_n^+, s_n^-. Then, setting $s_n = s_n^+ - s_n^-$ yields the desired sequence of simple functions. □

This theorem, combined with Proposition 2.4.5 gives us a nice characterization for Borel measurable functions: A function is Borel measurable if, and only if, it is the pointwise limit of simple functions. This will be a key result later on, when we make a kind of reasoning very common in measure theory: if we want to prove some property of some set of measurable functions, we first restrict ourselves to indicator functions, then we extend the property to simple functions, and finally we extend it to measurable functions via limits.

A first example of this kind of reasoning is the proof of following result, which we already showed for simple functions (this is Corollary 2.4.9):

Proposition 2.4.11. *The sum, product and division of measurable functions is measurable, provided it is defined.*

Proof. Let h_1 and h_2 be measurable functions, and approximate them by simple functions s_n^1, s_n^2 using Theorem 2.4.10. Then, wherever defined, results of arithmetic operations of measurable functions can be approximated in the following way:

- $s_n^1 + s_n^2 \to h_1 + h_2$
- $s_n^1 s_n^2 I_{\{h_1 \neq 0\}} I_{\{h_2 \neq 0\}} \to h_1 h_2$
- $\dfrac{s_n^1}{s_n^2 + (1/n) I_{\{s_n^2 = 0\}}} \to \dfrac{h_1}{h_2}$ □

Remark 2.4.12. *A consequence of this theorem is that any extended real-valued function is Borel measurable if, and only if, its negative and positive parts are.*

We finally have developed all the machinery required to define the Lebesgue integral and show some of its properties. Let $(\Omega, \mathscr{F}, \mu)$ be a measure space that will be fixed throughout the discussion.

Definition 2.4.13. *Let s be a simple function with domain Ω. Write*

$$s = \sum_i x_i I_{A_i}.$$

*We define the **Lebesgue integral** of s with respect to μ, and denote it by $\int_\Omega s \, d\mu$, $\int_\Omega s(\omega) \, d\mu$ or $\int_\Omega s(\omega) \, \mu(d\omega)$, as*

$$\int_\Omega s \, d\mu = \sum_i x_i \mu(A_i).$$

[10] The exposition of this proof is original. A much more *straight-to-the-point* proof is found in [6].

It is easy to check that this definition is well-posed in the sense that if s admits a different expression in terms of sums and products of indicator functions, then the sum on the right coincides.

In analysis, it is often very useful to be able to exchange symbols regarding limits and integrals or derivatives. In this case, we would like to be able to exchange signs while under the conditions of Theorem 2.4.10, that is, if $s_n \uparrow h$, where s_n are nonnegative, finite-valued, simple functions, then $\int_\Omega s_n \, d\mu. \uparrow \int_\Omega h \, d\mu$. This could be a way to define the integral for nonnegative Borel functions, but we would need to show that the limit does not depend on the sequence chosen. We can work around this by using suprema:

Definition 2.4.14. *Let h be a real-valued, Borel measurable function with domain Ω. If h is nonnegative, define*

$$\int_\Omega h \, d\mu = \sup\left\{\int_\Omega s \, d\mu \colon 0 \le s \le h, s \text{ simple}\right\}$$

For the general case, split $h = h^+ - h^-$ and set

$$\int_\Omega h \, d\mu = \int_\Omega h^+ \, d\mu - \int_\Omega h^- \, d\mu \text{ whenever the expression is not of the form } +\infty - \infty.$$

If it is of the form $+\infty - \infty$, we say the integral is not defined. Moreover, if $\int_\Omega h \, d\mu$ is finite we say that h is ***integrable****.*
Finally, if $B \in \mathscr{F}$, set $\int_B h \, d\mu = \int_\Omega h I_B \, d\mu$

Now that we have defined the integral for the general class of measurable functions, we need to show that it satisfies all the good properties an integral should satisfy.

Proposition 2.4.15. *Let g, h be extended real-valued, Borel measurable functions. Then,*

(i) If $\int_\Omega h \, d\mu$ exists, so does $\int_\Omega ch \, d\mu$ and $\int_\Omega ch \, d\mu = c\int_\Omega h \, d\mu$ for every $c \in \overline{\mathbb{R}}$.

(ii) The integral sign is monotonous. That is, if $g \le h$, then

$$\int_\Omega g \, d\mu \le \int_\Omega h \, d\mu$$

in the sense that if $\int_\Omega g \, d\mu$ exists and is greater that $-\infty$, then $\int_\Omega h \, d\mu$ exists; if $\int_\Omega h \, d\mu$ exists and is lesser than $+\infty$, then $\int_\Omega g \, d\mu$ exists; and whenever both integrals exist the inequality holds.

(iii) If $\int_\Omega h \, d\mu$ exists, then $\left|\int_\Omega h \, d\mu\right| \le \int_\Omega |h| \, d\mu$.

(iv) If h is nonnegative and $B \in \mathscr{F}$, then $\int_B h \, d\mu = \sup\left\{\int_B s \, d\mu \colon 0 \le s \le h, s \text{ simple}\right\}$

(v) If $\int_\Omega h \, d\mu$ exists, then so does $\int_B h \, d\mu$ for each $B \in \mathscr{F}$. If $\int_\Omega h \, d\mu$ is finite, so is $\int_B h \, d\mu$ for each $B \in \mathscr{F}$.

Proof. 1. The result clearly holds when h is simple. If $c=0$ it is also clearly true. If h is nonnegative and $c>0$,

$$\begin{aligned}\left\{\int_\Omega s\,d\mu\colon 0\le s\le ch, s\text{ simple}\right\} &= c\left\{\int_\Omega s/c\,d\mu\colon 0\le s\le ch, s\text{ simple}\right\}\\ &= c\left\{\int_\Omega s/c\,d\mu\colon 0\le s/c\le h, s/c\text{ simple}\right\},\\ &= c\left\{\int_\Omega s\,d\mu\colon 0\le s\le h, s\text{ simple}\right\}\end{aligned}$$

where in the second-to-last step we used the fact that s is simple if, and only if, s/c is simple. Taking suprema, $\int_\Omega ch\,d\mu = c\int_\Omega h\,d\mu$.

Now, if h is arbitrary, then for $c>0$, it holds that $(ch)^+ = ch^+$ and $(ch)^- = ch^-$. Hence, by what we just proved,

$$\int_\Omega ch\,d\mu = \int_\Omega ch^-\,d\mu + \int_\Omega ch^-\,d\mu = c\int_\Omega h^+\,d\mu + c\int_\Omega h^-\,d\mu = c\int_\Omega h\,d\mu.$$

If $c<0$, then $(ch)^+ = -ch^-$ and $(ch)^- = -ch^+$. Thus,

$$\int_\Omega ch\,d\mu = \int_\Omega -ch^-\,d\mu - \int_\Omega -ch^+\,d\mu = -c\int_\Omega h^-\,d\mu + c\int_\Omega h^+\,d\mu = c\int_\Omega h\,d\mu.$$

2. If g is nonnegative, then h is nonnegative too. The result follows immediately from the definition of the integral:

$$\left\{\int_\Omega s\,d\mu\colon 0\le s\le g\right\}\subseteq\left\{\int_\Omega s\,d\mu\colon 0\le s\le h\right\}.$$

Thus, $\int_\Omega g\,d\mu\le\int_\Omega h\,d\mu$. For the general case, note that $g\le h$ if, and only if, $g^+\le h^+$ and $h^-\le g^-$. Therefore,

$$\int_\Omega g^+\,d\mu\le\int_\Omega h^+\,d\mu \text{ and } \int_\Omega h^-\,d\mu\le\int_\Omega g^-\,d\mu.$$

From this, if $\int_\Omega g\,d\mu>-\infty$, then $\int_\Omega h^-\,d\mu\le\int_\Omega g^- <\infty$, and so $\int_\Omega h\,d\mu$ exists. The case where $\int_\Omega h\,d\mu<+\infty$ is similar. If both integrals exist,

$$\int_\Omega g\,d\mu = \int_\Omega g^+\,d\mu - \int_\Omega g^-\,d\mu \le \int_\Omega h^+\,d\mu - \int_\Omega h^-\,d\mu = \int_\Omega h\,d\mu.$$

3. This follows from the previous item and the fact that $-|h|\le h\le|h|$.

4. Note that $0\le s\le hI_B$ implies that $s=sI_B$, and $\int_\Omega s\,d\mu = \int_B s\,d\mu$. Therefore,

$$\begin{aligned}\left\{\int_\Omega s\,d\mu\colon 0\le s\le hI_B, s\text{ simple}\right\} &= \left\{\int_B s\,d\mu\colon 0\le sI_B\le hI_B, s\text{ simple}\right\}\\ &= \left\{\int_B s\,d\mu\colon 0\le s\le h, s\text{ simple}\right\}\end{aligned}.$$

Hence, the result follows taking suprema.

5. This follows from 2.4.15.(ii) and the facts that $(hI_B)^+ = h^+ I_B$ and $(hI_B)^- = h^- I_B$. $\square$

Remark 2.4.16. *Proposition 2.4.15.(iv) is interesting because it tells us that if we take some nonempty subset $A \in \mathscr{F}$ and regard it as a subspace of Ω (that is, consider the measure space $\left(A, \mathscr{F}_A, \mu|_{\mathscr{F}_A}\right)$, with $\mathscr{F}_A$ interpreted as $\mathscr{F}_A = \{B \cap A \colon B \in \mathscr{F}\}$), then the integral in this measure space coincides with the previously defined integral $\int_A h\, d\mu = \int_\Omega h I_A\, d\mu$.*

Now we have all the tools needed to prove a series of very powerful theorems concerning Lebesgue integration.

Lemma 2.4.17. *Let s be a nonnegative simple function defined on a measure space $\left(\Omega, \mathscr{F}, \mu\right)$. Then, the function $\lambda(B) = \int_B s\, d\mu$ is a measure on $\mathscr{F}$.*

Proof. If s is an indicator I_A, then $\int_B s\, d\mu = \mu(A \cap B)$, and in this case λ is clearly a measure because μ is. If we write s in its standard form $s = \sum_i^n x_i I_{A_i}$, then

$$\int_B s\, d\mu = \sum_i^n x_i \int_B I_{A_i}\, d\mu,$$

which clearly implies that λ is a measure in this case too. $\square$

Theorem 2.4.18 (Monotone Convergence Theorem). *Let $h_1, h_2 \dots$ be an increasing sequence of nonnegative Borel measurable functions converging to a pointwise limit h. Then, $\int_\Omega h_n\, d\mu \uparrow \int_\Omega h\, d\mu$.*

Proof. Since $h_n \le h$, by Proposition 2.4.15.(ii), the sequence of integrals $\int_\Omega h_n\, d\mu$ is increasing and bounded above by $\int_\Omega h\, d\mu$. Therefore, the limit $\lim_n \int_\Omega h_n\, d\mu$ exists and satisfies $\lim_n \int_\Omega h_n\, d\mu \le \int_\Omega h\, d\mu$. Let $k = \lim_n \int_\Omega h_n\, d\mu$.
Now let s be a simple function satisfying $0 \le s \le h$. Let $b \in (0,1)$, and define a sequence of sets $B_n = \{h_n \ge bs\}$[11]. It is clear that $B_n \uparrow \Omega$ and that $\int_{B_n} h_n\, d\mu \ge b \int_{B_n} s\, d\mu$. Therefore, by Proposition 2.4.15.(ii),

$$\int_\Omega h_n\, d\mu \ge \int_{B_n} h_n\, d\mu \ge b \int_{B_n} s\, d\mu.$$

Note that, by Lemma 2.4.17, the function $\lambda(B) = \int_B s\, d\mu$ is a measure. By taking limits when $n \to +\infty$ and following Proposition 2.2.11.(i), we have $k \ge b \int_\Omega s\, d\mu$. By taking limits when $b \to 1$, we have $k \ge \int_\Omega s\, d\mu$. By taking suprema for s, $k \ge \int_\Omega h\, d\mu$. $\square$

Theorem 2.4.19. *Let h be an extended real-valued, Borel measurable function such that $\int_\Omega h\, d\mu$ exists. Then, the function $\lambda(B) = \int_B h\, d\mu$ is countably additive. In particular, if $h \ge 0$, then λ is a measure.*

[11] This set is measurable: write s in standard form and express B_n as an intersection of measurable sets.

Proof. First suppose h nonnegative. Use Theorem 2.4.10 to obtain a sequence of simple functions s_m with $s_m \uparrow h$. Let $B_1, B_2, \ldots$ be a sequence of disjoint measurable sets, and let $B = \bigcup_n B_n$. Note that $s_m I_B \uparrow hI_B$. By the Monotone Convergence Theorem, $\int_B s_m \, d\mu \uparrow_m \int_B h \, d\mu$. Since, for every m, the function defined by $\lambda_m(A) = \int_A s_m \, d\mu$ is a measure by Lemma 2.4.17, we have $\int_B s_m \, d\mu = \sum_n \int_{B_n} s_m \, d\mu$. Define the sequence $a_{nm} = \sum_{k=1}^n \int_{B_k} s_m \, d\mu$. Then, by what we have just seen,

$$\lambda(B) = \lim_m \lim_n a_{nm}$$

Since a_{nm} is increasing with respect to both indices, we can apply Lemma A.1.1, and thus

$$\lambda(B) = \lim_n \lim_m a_{nm} = \lim_n \sum_{k=1}^n \lim_m \int_{B_k} s_m \, d\mu$$

However, for every n, $s_m I_{B_n} \uparrow_m hI_{B_n}$ which implies that $\lim_m \int_{B_n} s_m \, d\mu = \int_{B_n} h \, d\mu$. Therefore, $\lambda(B) = \sum_n \int_{B_n} h \, d\mu = \sum_n \lambda(B_n)$, as we wanted to see.
For the general case, split $h = h^+ - h^-$. Apply the result proved so far to h^+ and h^- to obtain two measures, λ^+ and λ^-. Since $\int_\Omega h \, d\mu$ exists, at least one of $\int_\Omega h^+ \, d\mu$ and $\int_\Omega h^- \, d\mu$ is finite, and therefore at least one of λ^+ and λ^- is finite, which ensures that λ is well defined and σ-additive. □

With this last result,[12] we can finally prove that the integral is additive:

Proposition 2.4.20. *Let f, g be Borel measurable functions, and assume that $f+g$ is well-defined. If $\int_\Omega f \, d\mu$ and $\int_\Omega g \, d\mu$ exist and their sum is well-defined, then the integral $\int_\Omega f+g \, d\mu$ exists and*

$$\int_\Omega f+g \, d\mu = \int_\Omega f \, d\mu + \int_\Omega g \, d\mu.$$

Proof. If f and g are nonnegative simple functions, the result follows easily from the definition of the integral. If f and g are nonnegative, we can approximate them by increasing sequences of nonnegative simple functions via Theorem 2.4.10: $s_n^1 \uparrow f, s_n^2 \uparrow g$. Thus, by the Monotone Convergence Theorem,

$$\int_\Omega s_n^1 + s_n^2 \, d\mu = \int_\Omega s_n^1 \, d\mu + \int_\Omega s_n^2 \, d\mu \uparrow \int_\Omega f+g \, d\mu = \int_\Omega f \, d\mu + \int_\Omega g \, d\mu.$$

The proof of the general case is a casewise proof (depending on the signs of f, g and $f+g$; and splitting Ω into the subsets where each combination of signs takes place) with little interesting ideas - except from the one exposed - and is omitted. □

Corollary 2.4.21. *(i) If $h_1, h_2, \ldots$ are nonnegative Borel measurable, then*

$$\sum_{n=1}^{+\infty} \int_\Omega h_n \, d\mu = \int_\Omega \left(\sum_{n=1}^{+\infty} h_n \right) d\mu$$

(ii) If h is Borel measurable, h is integrable if, and only if, $|h|$ is.

[12] In [6], Theorem 2.4.19 is proved without employing the previous two results used here. However, the proof is, admittedly, rather technical and difficult. The approach followed here is original and (hopefully) a more comprehensible one.

(iii) If g and h are Borel measurable, $|g| \leq h$ and h is integrable, then g is integrable.

Proof. 1. Direct from the additivity of the integral, the Monotone Convergence Theorem and the fact that $\sum_{n=1}^{N} h_n \uparrow \sum_{n=1}^{+\infty} h_n$ when $N \to +\infty$.

2. If $|h| = h^+ + h^-$ is integrable, it follows from additivity that so are h^+ and h^-. Thus, h is integrable. If h is integrable, then so are by h^+ and h^- by the definition of integral. Thus, h is integrable by additivity.

3. $|g|$ is integrable because of monotonicity, and by 2.4.21.(ii), g is integrable. □

Definition 2.4.22. *A condition is said to hold almost everywhere with respect to a measure μ (written μ-a.e., a.e. [μ] or simply a.e., if there is no confusion respect the measure of integration) whenever there exists some $B \in \mathscr{F}$ where the property is satisfied and such that $\mu(B^c) = 0$.*

An important result is that, from the integration point of view, two functions that coincide almost everywhere are identical. This is captured in the following proposition:

Proposition 2.4.23. *Let f, g and h be Borel measurable functions.*

(i) If $f = 0$ μ-a.e., then $\int_\Omega f \, d\mu = 0$.

(ii) If $g \leq h$ μ-a.e., then $\int_\Omega g \, d\mu \leq \int_\Omega h \, d\mu$ in the sense of Proposition 2.4.15.(ii).

(iii) If $g = h$ μ-a.e., then $\int_\Omega g \, d\mu = \int_\Omega h \, d\mu$, in the sense that one exists if, and only if, the other one does, and they are equal.

Proof. 1. If s is simple and nonnegative, we can write s in standard form as

$$s = \sum_i x_i I_{A_i}.$$

If we define $N = \{s \neq 0\}$, then $x_i \neq 0$ implies $A_i \subseteq N$, whence $\mu(A_i) \leq \mu(N) = 0$.

If $f \geq 0$ and s is a simple function such that $0 \leq s \leq f$, then $s = 0$ a.e., and thus $\int_\Omega s \, d\mu = 0$. It follows that $\int_\Omega f \, d\mu = 0$.

For the general case, $f = 0$ a.e. implies $|f| = 0$ a.e. Since f^+ and f^- are both nonnegative and bounded above by $|f|$, they are both null a.e. Thus, $\int_\Omega f^+ \, d\mu = \int_\Omega f^- \, d\mu = 0$.

2. Define $B = \{g > h\}$[13]. Let $A = B^c$. $g = gI_A + gI_B$ and $h = hI_A + hI_B$. Since gI_B and hI_B, by 2.4.23.(i) their integrals are 0. Then, by Proposition 2.4.15.(ii),

$$\int_\Omega g \, d\mu = \int_A g \, d\mu \leq \int_A h \, d\mu = \int_\Omega h \, d\mu$$

[13] It is not immediate that B is measurable. One way to see it is by writing $B = \{h < +\infty\} \cap \{g > -\infty\} \cap \bigcap_{q \in \mathbb{Q}} (\{g > q\} \cap \{q > h\})$.

3. Define $B = \{g \neq h\} = \{g > h\} \cup \{h > g\}$, and $A = B^c$. Therefore,

$$\int_\Omega g\, d\mu = \int_A g\, d\mu = \int_A h\, d\mu = \int_\Omega h\, d\mu. \qquad \square$$

Proposition 2.4.24. *(i) If h is integrable with respect to μ, then it is finite μ-a.e.*

(ii) If $h \geq 0$ and $\int_\Omega h\, d\mu = 0$, then $h = 0$ a.e.

Proof. 1. Let $B = \{h \in \mathbb{R}\}$ and $A = B^c$. Then, if $\mu(A) > 0$, we would have

$$\int_\Omega |h|\, d\mu \geq \int_A |h| = \infty \cdot \mu(A) = \infty,$$

which is a contradiction.

2. Define $B = \{h > 0\}$ and $B_n = \left\{h > \frac{1}{n}\right\}$. Note that $B_n \uparrow B$. Therefore, $\mu(B_n) \uparrow \mu(B)$ and $\int_{B_n} h\, d\mu \uparrow \int_B h\, d\mu$ by Theorem 2.4.19 and Proposition 2.2.11.(i). Note that, since $0 \leq \int_{B_n} h\, d\mu \leq \int_B h\, d\mu = 0$, we have $\int_{B_n} h\, d\mu = 0$. However, in B_n, $h \geq 1/n$, whence $\int_{B_n} h\, d\mu \geq \mu(B_n)/n$. One deduces that $\mu(B_n) = 0$ for all n. Therefore, $\mu(B) = 0$. $\square$

As we said, differences between functions in null sets are irrelevant regarding integration. The Monotone Convergence Theorem can be extended in order to account for this, and additionally, the nonnegativity hypothesis can be greatly relaxed.

Theorem 2.4.25 (Extended Monotone Convergence Theorem). *Let $g_1, g_2 \ldots$ be a sequence of Borel measurable functions.*

(i) Suppose that $\int_\Omega g_1\, d\mu > -\infty$ and $g_n \uparrow g$ a.e., that is, we have $g = \lim_n g_n$ a.e. and, for every n, $g_n \leq g_{n+1}$ a.e. Then, the integrals $\int_\Omega g\, d\mu$ and $\int_\Omega g_n\, d\mu$ exist for all n, and

$$\int_\Omega g_n\, d\mu \uparrow \int_\Omega g\, d\mu.$$

(ii) Suppose that $\int_\Omega g_1\, d\mu < +\infty$ and $g_n \downarrow g$ a.e., that is, we have $g = \lim_n g_n$ a.e. and, for every n, $g_n \geq g_{n+1}$ a.e. Then, the integrals $\int_\Omega g\, d\mu$ and $\int_\Omega g_n\, d\mu$ exist for all n, and

$$\int_\Omega g_n\, d\mu \downarrow \int_\Omega g\, d\mu.$$

Proof. 1. Let $P_n = \{g_n > g_{n+1}\}$, and $P = \bigcup_n P_n$. Note that $\mu(P) = 0$. Let L be the set where $g = \lim_n g_n$, and $G = P \cup L^c$. Note that $\mu(G) = 0$ too. Define $\overline{g} = g I_{G^c}$ and a sequence of functions $\overline{g}_n = g_n I_{G^c}$, so that $\overline{g}_n \uparrow \overline{g}$, $\int_\Omega \overline{g}\, d\mu = \int_\Omega g\, d\mu$ and $\int_\Omega \overline{g}_n\, d\mu = \int_\Omega g_n\, d\mu$ for every n. Since $\int_\Omega \overline{g}_1\, d\mu > -\infty$, it must be that $\overline{g_1}^-$ is integrable, and thus finite almost everywhere. Let A be the set where $\overline{g_1}^-$ is finite. Note that, since the sequence $\{\overline{g}_n\}_{n\in\mathbb{N}}$ is increasing, then $\{\overline{g}_n^-\}_{n\in\mathbb{N}}$ is decreasing. Therefore, every $\overline{g}_n^-$ is integrable and finite on A. This also tells us that $\int_\Omega \overline{g}_n\, d\mu$ exists for all n. Define $h_n = I_A \cdot \overline{g}_n + I_A \overline{g_1}^-$, so that $\int_\Omega h_n\, d\mu = \int_\Omega g_n\, d\mu + \int_\Omega \overline{g_1}^-\, d\mu$.

Then, the functions h_n are nonnegative and increase to $I_A \cdot \overline{g} + I_A \overline{g}_1^-$. By the Monotone Convergence Theorem,

$$\int_\Omega g_n \, d\mu + \int_\Omega \overline{g}_1 \, d\mu \uparrow \int_\Omega g \, d\mu + \int_\Omega \overline{g}_1^- \, d\mu$$

The proof is completed by noting that $\int_\Omega \overline{g}_1 \, d\mu$ is finite, and can therefore be substracted from both sides of the expression above.

2. Apply the previous section to $\{-g_n\}_{n\in\mathbb{N}}$ and $-g$.[14] □

Note that, if $\{g_n\}_{n\in\mathbb{Z}^+}$ is a sequence of Borel measurable functions, then $g = \sup_n f_n$ and $h = \inf_n f_n$ defined pointwise (that is, $g(\omega) = \sup_n f_n(\omega)$, and similarly for h) are Borel measurable too: for all $c \in \overline{\mathbb{R}}$,

$$\{g \le c\} = \bigcap_n \{f_n \le c\} \in \mathscr{F}.$$

Therefore, if we define them pointwise as well, $\liminf_n f_n$ and $\limsup_n f_n$ are measurable too. We can now present a very important theorem regarding these functions:

Theorem 2.4.26 (Fatou's Lemma). *Let $f_1, f_2, \dots, f$ be Borel measurable. Then,*

(i) If $f_n \ge f$ for all n, where $\int_\Omega f \, d\mu > -\infty$, then the integral $\int_\Omega \liminf_n f_n \, d\mu$ exists and

$$\liminf_n \int_\Omega f_n \, d\mu \ge \int_\Omega \liminf_n f_n \, d\mu.$$

(ii) If $f_n \le f$ for all n, where $\int_\Omega f \, d\mu < +\infty$, then the integral $\int_\Omega \limsup_n f_n \, d\mu$ exists and

$$\limsup_n \int_\Omega f_n \, d\mu \le \int_\Omega \limsup_n f_n \, d\mu.$$

Proof. 1. Define $g_n = \inf_{k\ge n} f_k$ and $g = \sup_n g_n = \liminf_n f_n$. Then, $g_n \uparrow g$ and $g_1 \ge f$, whence $\int_\Omega g_1 \, d\mu$ exists and is greater that $-\infty$. By Theorem 2.4.25, $\int_\Omega g_n \, d\mu \uparrow \int_\Omega g \, d\mu$. Moreover, since $g_n \le f_k$ for all $k \ge n$, we have $\int_\Omega g_n \, d\mu \le \int_\Omega f_k \, d\mu$ for all $k \ge n$, so $\int_\Omega g_n \, d\mu \le \inf_{k\ge n} \int_\Omega f_k \, d\mu$. Therefore, we have

$$\int_\Omega \liminf_n f_n \, d\mu = \int_\Omega g \, d\mu = \sup_n \int_\Omega g_n \, d\mu \ge \sup_n \inf_{k\ge n} \int_\Omega f_k \, d\mu = \liminf_n \int_\Omega f_n \, d\mu$$

2. Note that $-f_1, -f_2, \dots, -f$ satisfy all the hypotheses of the last item. Thus,

$$\liminf_n \int_\Omega -f_n \, d\mu \ge \int_\Omega \liminf_n -f_n \, d\mu.$$

The result is obtained by multiplying the inequality by -1. □

The following result can be obtained as a simple corollary of Fatou's Lemma. It is one of the most important theorems in analysis regarding integration:

[14] In this theorem, hypotheses were simplified with respect to [6], and the proof is (somewhat) original.

Theorem 2.4.27 (Dominated Convergence Theorem). *If $f_1, f_2, \ldots, f, g$ are Borel measurable functions, g is μ-integrable, $|f_n| \le g$ and $f_n \to f$ μ-a.e., then f is μ-integrable and*

$$\int_\Omega f\, d\mu = \lim_n \int_\Omega f_n\, d\mu.$$

Proof. By taking limits on the inequality $|f_n| \le g$, one deduces that $|f| \le g$ μ-a.e., hence f is integrable. Furthermore, $-g \le f_n \le g$. Since the sequence of f_n converges to f almost everywhere, we have $\liminf_n f_n = \limsup_n f_n = f$ μ-a.e. Thus, $\int_\Omega \liminf_n f_n\, d\mu = \int_\Omega \limsup_n f_n\, d\mu = \int_\Omega f\, d\mu$. By Fatou's lemma,

$$\liminf_n \int_\Omega f_n\, d\mu \ge \int_\Omega f\, d\mu \ge \limsup_n \int_\Omega f_n\, d\mu.$$

It follows that the limit $\lim_n \int_\Omega f_n\, d\mu$ exists and is equal to $\int_\Omega f\, d\mu$. □

Theorem 2.4.28. *If μ is σ-finite on $\mathscr{F}$, g and h are Borel measurable, $\int_\Omega g\, d\mu$ and $\int_\Omega h\, d\mu$ exist, and $\int_A g\, d\mu \le \int_A h\, d\mu$ for all $A \in \mathscr{F}$, then $g \le h$ μ-a.e.*

Proof. Decompose Ω into countably many subsets with finite measure A_n. Regard each A_n as a finite measure space. Then, it suffices to show that $g \le h$ μ-a.e. in A_n for all $n \in \mathbb{Z}^+$. This allows us to suppose, without loss of generality, that μ is finite.
Let F be the set where h is finite. Define the sets $B = \{g > h\} \cap F$ and, for each $n \in \mathbb{Z}^+$, $B_n = \{g \ge h + \frac{1}{n}\} \cap \{|h| \le n\} \cap F$, so that $B_n \uparrow B$. Now note that, on one side, $\int_{B_n} g\, d\mu \le \int_{B_n} h\, d\mu \le n\mu(B_n) < +\infty$ (because $\int_A g\, d\mu \le \int_A h\, d\mu$ for all A, and $h \le n$ on B_n), and on the other,

$$\int_{B_n} g\, d\mu \ge \int_{B_n} h\, d\mu + \mu(B_n)/n \ge \int_{B_n} g\, d\mu + \mu(B_n)/n.$$

This implies that $\mu(B_n) = 0$. Since $\mu(B_n) \uparrow \mu(B)$, it follows that $\mu(B) = 0$. Now let us consider $F^c = \{h = -\infty\} \cup \{h = +\infty\}$. Clearly, $g \le h$ on $\{h = +\infty\}$. Define $C = \{h = -\infty\} \cap \{g > h\}$, and $C_n = C \cap \{g \ge -n\}$. Therefore, $C_n \uparrow C$, and

$$-\infty \cdot \mu(C_n) = \int_{C_n} h\, d\mu \ge \int_{C_n} g\, d\mu \ge -n\mu(C_n),$$

hence $\mu(C_n) = 0$ ($\mu(C_n) = +\infty$ is impossible because μ is finite). Since $\mu(C_n) \uparrow \mu(C)$, we have $\mu(C) = 0$. □

Corollary 2.4.29. *If μ is σ-finite on $\mathscr{F}$, g and h are Borel measurable, $\int_\Omega g\, d\mu$ and $\int_\Omega h\, d\mu$ exist, and $\int_A g\, d\mu = \int_A h\, d\mu$ for all $A \in \mathscr{F}$, then $g = h$ μ-a.e.*

We end the section with two useful theorems. The fisrt is a very general form of a change of variables formula:

Theorem 2.4.30 (Image Measure Theorem). *Let $(\Omega_1, \mathscr{F}_1)$, $(\Omega_2, \mathscr{F}_2)$ be measurable spaces and μ_1 a measure on $\mathscr{F}_1$. Let $T \colon \Omega_1 \to \Omega_2$ be a measurable mapping.*
*Define a measure μ_2 on $\mathscr{F}_2$ by $\mu_2(A) = \mu_1(T^{-1}(A))$. The measure μ_2 is often called the **push-forward measure** or **image measure** of μ_1 by T and denoted by $T_*(\mu_1)$ or $\mu_1 \circ T^{-1}$.*

Then, for every Borel measurable function $f : \Omega_2 \to \overline{\mathbb{R}}$ and every $A \in \mathscr{F}_2$, one has

$$\int_A f\, d\mu_2 = \int_{T^{-1}(A)} \left(f \circ T\right) d\mu_1,$$

in the sense that if one integral exists, so does the other, and the two are equal.

Proof. First suppose that f is an indicator I_B. Then, $f \circ T = I_{T^{-1}(B)}$, and thus the desired formula becomes

$$\int_{\Omega_2} I_{A \cap B}\, d\mu_2 = \int_{\Omega_1} I_{T^{-1}(A) \cap T^{-1}(B)}\, d\mu_1,$$

which is equivalent to $\mu_2(A \cap B) = \mu_1(T^{-1}(A \cap B))$, and this is true by definition.
If f is a nonnegative simple function $\sum_{i=1}^n x_i I_{B_i}$, then

$$\int_A f\, d\mu_2 = \sum_{i=1}^n x_i \int_A I_{B_I}\, d\mu_2 = \sum_{i=1}^n x_i \int_{T^{-1}(A)} \left(I_{B_i} \circ T\right) d\mu_1 = \int_{T^{-1}(A)} \left(f \circ T\right) d\mu_1.$$

If f is a nonnegative Borel measurable function, take $s_1, s_2, \dots$ nonnegative simple functions increasing to f. Thus, $s_1 \circ T, s_2 \circ T, \dots$ are nonnegative simple functions increasing to $f \circ T$. The Monotone Convergence Theorem yields the result.
Finally, if $f = f^+ - f^-$, the result proved so far yields the desired formula for f^+ and f^-, since $(f \circ T)^+ = f^+ \circ T$ and $(f \circ T)^- = f^- \circ T$. If, say, $\int_A f^+ d\mu_2$ is finite, so is $\int_{T^{-1}(A)} \left(f^+ \circ T\right) d\mu_1$ and additivity implies the result for f. □

The second result is quite simple, both to state and prove, but it is a very useful theorem in analysis.

Theorem 2.4.31 (Borel-Cantelli Lemma). *Let μ be a measure on a σ-field $\mathscr{F}$. Let $A_1, A_2, \dots$ be a sequence of measurable sets. Then*

$$\sum_n \mu(A_n) < \infty \quad \text{implies} \quad \mu\left(\limsup_n A_n\right) = 0.$$

Proof. For every n, we have

$$\mu\left(\bigcup_{k \geq n} A_k\right) \leq \sum_{k \geq n} \mu(A_k).$$

It follows, by Proposition 2.2.11.(ii) that $\mu(\limsup_n A_n) \leq 0$, which in turn yields the desired result. □

CHAPTER 3

ADVANCED RESULTS IN MEASURE THEORY

In the previous part of the text, we developed the *basics* of Measure Theory. Many possibilities open now: for instance, it is possible to prove the famous Radon-Nikodým Theorem with the theory developed so far. This is done in great detail in Appendix B.

In this text, however, we will focus on developing all the theory necessary to prove the Kolmogorov Extension Theorem. This is what will be done in this chapter.

3.1 Daniell Theory

The study of linear functionals[1] is a very fruitful topic in Analysis. Weak topologies are based on the concept of continuity of linear functionals in a given Banach space, and they provide a solid basis for many useful theorems regarding convergence in function spaces [1].

The integral, in particular, is itself a linear functional from the space of integrable functions on a given measure space to $\mathbb{R}$. As we will see, under appropiate hypotheses it is possible to construct a measure so that a given linear functional can be expressed as the integral over that measure. The approach followed to do this will be that of Daniell Theory. But first, we need some basic concepts and tools:

Definition 3.1.1. *Let $\mathscr{D}$ be a class of subsets of some nonempty set Ω. We say that $\mathscr{D}$ is a* ***Dynkin system*** *(or* ***D-system*** *for short) if*

(i) $\Omega \in \mathscr{D}$.

(ii) *$\mathscr{D}$ is closed under set differences; that is, if $A \subseteq B$ with $A, B \in \mathscr{D}$, then $A \setminus B \in \mathscr{D}$.*

(iii) *$\mathscr{D}$ is closed under increasing sequences; that is, if $A_1, \ldots, A_n, \ldots$ is a sequence of sets in $\mathscr{D}$ and $A_n \uparrow A$, then $A \in \mathscr{D}$.*

[1] A *functional* is a function from a given vector space of real-valued functions to $\mathbb{R}$.

By conditions 3.1.1.(i) and 3.1.1.(ii), $\mathscr{D}$ is closed under complementation. By conditions 3.1.1.(ii) and 3.1.1.(iii), $\mathscr{D}$ is a monotone class. If $\mathscr{D}$ is closed under finite unions (or intersections), then it is a σ-field.

The arbitrary intersection of D-systems is a D-system too. This guarantees the existence of generated D-systems. If $\mathscr{S}$ is a class of subsets of some nonempty set Ω, we will denote its generated D-system - that is, the smallest D-system containing $\mathscr{S}$ - by $\mathscr{D}(\mathscr{S})$.

Our interest on D-systems is motivated by the following theorem. It is analogous to the Monotone Class Theorem.

Theorem 3.1.2 (Dynkin System Theorem). *Let $\mathscr{S}$ be a class of subsets of Ω. If $\mathscr{S}$ is closed under finite intersection, then $\mathscr{D}(\mathscr{S}) = \sigma(\mathscr{S})$. In particular, if $\mathscr{D}$ is a Dynkin system and $\mathscr{S} \subseteq \mathscr{D}$, then $\sigma(\mathscr{S}) \subseteq \mathscr{D}$.*

Proof. Let $\mathscr{D}_0 = \mathscr{D}(\mathscr{S})$ and $\mathscr{F} = \sigma(\mathscr{S})$. Define $\mathscr{V} = \{A \in \mathscr{D}_0 \,|\, A \cap B \in \mathscr{D}_0 \text{ for every } B \in \mathscr{S}\}$. Now, $\mathscr{S} \subseteq \mathscr{V}$ since $\mathscr{S}$ is closed under intersection and a subset of $\mathscr{D}_0$. Also, it is easy to check that $\mathscr{V}$ is a D-system because $\mathscr{D}_0$ is. Thus, $\mathscr{V} = \mathscr{D}_0$. Now define $\mathscr{V}' = \{A \in \mathscr{D}_0 \,|\, A \cap B \in \mathscr{D}_0 \text{ for every } B \in \mathscr{D}_0\}$. Again, $\mathscr{V}'$ is a D-system and, clearly, $\mathscr{S} \subseteq \mathscr{V} \subseteq \mathscr{V}'$; hence $\mathscr{V}' = \mathscr{D}_0$.

It follows that $\mathscr{D}_0$ is closed under finite intersection, hence a σ-field. Thus, $\mathscr{F} \subseteq \mathscr{D}_0$. The other inclusion is immediate because every σ-field is a D-system. Finally, $\mathscr{F} = \mathscr{D}_0 \subseteq \mathscr{D}$. □

In the Monotone Class Theorem, a weaker hypothesis is imposed to the generating set (in this theorem, $\mathscr{S}$ need not be a field), but a stronger hypothesis is imposed to the structure (a monotone class need not be a D-system).

Corollary 3.1.3. *Let $\mathscr{S}$ be a class of subsets of Ω that is closed under finite intersection and such that $\Omega \in \mathscr{S}$. If μ_1 and μ_2 are finite measures on $\sigma(\mathscr{S})$ that agree on $\mathscr{S}$, then $\mu_1 = \mu_2$ on $\sigma(\mathscr{S})$.*

Proof. Let $\mathscr{D}$ be the class of sets of $\sigma(\mathscr{S})$ where μ_1 and μ_2 agree. Then, $\mathscr{D}$ is a D-system:

- Condition 3.1.1.(i) follows from the fact that $\Omega \in \mathscr{S}$.
- Condition 3.1.1.(ii) follows from additivity.
- Condition 3.1.1.(iii) follows from Proposition 2.2.7.(iii).

Hence, $\sigma(\mathscr{S}) = \mathscr{D}(\mathscr{S}) \subseteq \mathscr{D}$. The reciprocal inclusion is true by definition. □

Corollary 3.1.4. *Let $\mathscr{S}$ be a class of subsets of Ω that is closed under finite intersection and such that $\Omega \in \mathscr{S}$. Let H be a vector space of real-valued functions on Ω, such that $I_A \in H$ for each $A \in \mathscr{S}$. Suppose that for every increasing sequence of nonnegative functions $f_1, f_2 \dots$ with a bounded limit (that is, $f_n \uparrow f$ and there exists some $M \in \mathbb{R}^+$ such that $|f| \leq M$), the limit function f belongs to H.*
Then, $I_A \in H$ for every $A \in \sigma(\mathscr{S})$.

Proof. Let $\mathscr{D}_0 = \mathscr{D}(\mathscr{S})$ and $\mathscr{D} = \{A \in \mathscr{D}_0 \,|\, I_A \in H\}$. Then, $\mathscr{D}$ is a D-system, because:

- $\Omega \in \mathscr{D}$ since $\Omega \in \mathscr{S}$.

- If $A \subseteq B$, $A, B \in \mathscr{D}$, then $I_{B \setminus A} = I_B - I_A \in H$.
- If $A_n \uparrow A$, $A_n \in \mathscr{D}$, then $I_{A_n} \uparrow I_A \in H$.

Thus, $\mathscr{D} = \mathscr{D}_0 = \sigma(\mathscr{S})$, completing the proof. □

Definition 3.1.5. *Let L be a vector space of real-valued functions on a set Ω. We will say that L is closed under the* ***lattice operations*** *if $\max(f, g) \in L$ and $\min(f, g) \in L$ for every two functions $f, g \in L$.*
If $E\colon L \to \mathbb{R}$ is a linear functional, we say that it is ***positive*** *if $f \geq 0$ implies $E(f) \geq 0$. From this, it follows that E is* ***monotone****; that is, $f \geq g$ implies $E(f) \geq E(g)$. Additionally, we will say that E is a* ***Daniell integral*** *if $f_n \uparrow f$, $f_n \geq 0$ implies $E(f_n) \uparrow E(f)$ and $f_n \downarrow 0$ implies $E(f_n) \downarrow E(0) = 0$*[2]*.*
If H is any class of functions from Ω to $\overline{\mathbb{R}}$, H^+ will denote the class of nonnegative functions in H, $\{f \in H \,|\, f \geq 0\}$. The collection of functions $f\colon \Omega \to \overline{\mathbb{R}}$ such that there exists a sequence f_n in L^+ with $f_n \uparrow f$ will be denoted by L'.
If H is as above, the σ-field ***generated*** *by H, denoted by $\sigma(H)$, is defined as the smallest σ-field making all functions in H Borel measurable; namely, $\sigma(H) = \sigma(\mathscr{A})$, where $\mathscr{A} = \left\{f^{-1}(B) \,\middle|\, f \in H, B \in \mathscr{B}\left(\overline{\mathbb{R}}\right)\right\}$.*

During the rest of the section, L will be a vector space as above, and E will be a Daniell integral on L.
A great part of this section will follow a structure very similar to that of Section 2.3. We begin by extending E to L':

Lemma 3.1.6. *Let $\{f_n\}$ and $\{g_n\}$ be sequences in L increasing to respective limits f and g, with $f \leq g$. Then,*

$$\lim_n E(f_n) \leq \lim_n E(g_n).$$

Hence, E may be extended to L' as $E(\lim_n h_n) = \lim_n E(h_n)$.

Proof. First, note that both limits exist (they may be $+\infty$) because we have increasing sequences of real numbers. Now, $\min(f_m, g_n) \uparrow_n \min(f_m, g) = f_m$. Thus, $E(f_m) = \lim_n E(\min(f_m, g_n)) \leq \lim_n E(g_n)$. Take limits in m to complete the proof. □

We now study this extension to L':

Lemma 3.1.7. *The extension of E to L' has the following properties:*

(i) $0 \leq E(f) \leq +\infty$ *for all $f \in L'$.*

(ii) If $f, g \in L'$ and $f \leq g$, then $E(f) \leq E(g)$.

(iii) If $f \in L'$ and c is a nonnegative real number, then $cf \in L'$ and $E(cf) = cE(f)$.

(iv) If $f, g \in L'$, then $f + g$, $\min(f, g)$ and $\max(f, g)$ all are in L', and

$$E(f + g) = E(f) + E(g) = E(\min(f, g)) + E(\max(f, g)).$$

[2] In practice, if we want to see that a given linear functional is a Daniell integral, it suffices to show that E is positive and that $f_n \downarrow 0$ implies $E(f_n) \downarrow 0$.

(v) *If* $f_n \in L'$ *and* $f_n \uparrow f$, *then* $E(f_n) \uparrow E(f)$.

Proof. Items (i)-(iii) are immediate, either by definition or by Lemma 3.1.6.
To see Lemma 3.1.7.(iv), simply take sequences $f_n \uparrow f$, $g_n \uparrow g$ in L', so that $\min(f_n, g_n) \uparrow \min(f, g) \in L'$, $\max(f_n, g_n) \uparrow \max(f, g) \in L'$ and $f_n + g_n \uparrow f + g \in L'$. Additionally, $E(f + g) = \lim_n E(f_n + g_n) = \lim_n E(f_n) + E(g_n) = E(f) + E(g)$. The last equality follows from linearity and the fact that $f + g = \max(f, g) + \min(f, g)$.
To see 3.1.7.(v), for each $n \in \mathbb{Z}^+$ consider a sequence f_{nm} in L^+ such that $f_{nm} \uparrow_m f_n$. Define $g_m = \max(f_{1m}, \dots, f_{mm}) \in L^+$, so that

$$f_{nm} \le g_m \le f_m \tag{3.1}$$

for $n \le m$. Take $m \to +\infty$ to see that $f_n \le \lim_m g_m \le f$, and then $n \to +\infty$ to obtain $g_m \uparrow f$; hence $E(g_m) \uparrow E(f)$. Now apply E in equation (3.1) to obtain $E(f_{nm}) \le E(g_m) \le E(f_m)$. Let $m \to +\infty$ to obtain $E(f_n) \le E(f) \le \lim_m E(f_m)$. Now let $n \to +\infty$ to obtain the desired result. □

We now begin the construction of a σ-field and a measure derived from E. Henceforth, we assume that all constant functions belong to L. We can rescale E so that $E(1) = 1$ (hence, $E(c) = c$ for all $c \in \mathbb{R}$).

Lemma 3.1.8. *Let* $\mathscr{C}$ *be the class of subsets* $G \subseteq \Omega$ *such that* $I_G \in L'$ *and define* $\mu(G) = E(I_G)$. *Then,* $\mathscr{C}$ *and* μ *satisfy all four conditions of Lemma 2.3.2; namely:*

(i) $\emptyset, \Omega \in \mathscr{C}$, $\mu(\emptyset) = 0$, $\mu(\Omega) = 1$ *and* $0 \le \mu(A) \le 1$ *for all* $A \in \mathscr{C}$

(ii) *If* $G_1, G_2 \in \mathscr{C}$, *then* $G_1 \cup G_2, G_1 \cap G_2 \in \mathscr{C}$ *and* $\mu(G_1 \cup G_2) + \mu(G_1 \cap G_2) = \mu(G_1) + \mu(G_2)$.

(iii) *If* $G_1, G_2 \in \mathscr{C}$ *and* $G_1 \subseteq G_2$, *then* $\mu(G_1) \le \mu(G_2)$.

(iv) *If* $G_n \in \mathscr{C}$, *and* $G_n \uparrow G$, *then* $G \in \mathscr{C}$ *and* $\mu(G_n) \to \mu(G)$.

Hence, by Lemma 2.3.3 and Theorem 2.3.4, the mapping $\mu^*(A) = \inf\{\mu(G) \,|\, G \in \mathscr{C}, A \subseteq G\}$ *is a probability measure on the* σ*-field* $\mathscr{H} = \{H \subseteq \Omega \,\big|\, \mu^*(H) + \mu^*(H^c) = 1\}$ *such that* $\mu \equiv \mu^*$ *on* $\mathscr{C}$.

Proof.
1. Since L contains all constant functions, $I_\emptyset = 0$ and $I_\Omega = 1$ are in L. Additionally, $E(c) = c$; in particular, $E(I_\emptyset) = 0$ and $E(I_\Omega) = 1$. The rest follows from E being monotone.
2. Direct consequence of Lemma 3.1.7.(iv), taking $f = I_{G_1}$ and $g = I_{G_2}$. Simply note that $\min(I_{G_1}, I_{G_2}) = I_{G_1 \cap G_2}$ and $\max(I_{G_1}, I_{G_2}) = I_{G_1 \cup G_2}$.
3. Immediate by the monotonicity of E.
4. Consequence of Lemma 3.1.7.(v), taking $f_n = I_{G_n}$; note that $f = I_G$. □

Just as in Section 2.3, it is true that $\sigma(\mathscr{C}) \subseteq \mathscr{H}$. However, the proof of this is a little harder this time, and we will need some previous results.
First, we investigate Borel measurability of functions on L' relative to the σ-field $\sigma(\mathscr{C})$.

Lemma 3.1.9. *If $f \in L'$ and $a \in \mathbb{R}$, then the set $Z = \{f(\omega) > a\}$ belongs to the class $\mathscr{C}$. Therefore, f is Borel measurable with respect to $\sigma(\mathscr{C})$.*

Proof. Let f_n be a sequence in L^+ such that $f_n \uparrow f$. For any $a \in \mathbb{R}$, $(f_n - a)^+ = \max(f_n - a, 0) \in L^+$, so that $(f - a)^+ = \lim_n (f_n - a)^+ \in L'$. Then, by Lemma 3.1.7.(iii), for every $k \in \mathbb{Z}^+$, $k(f - a)^+ \in L'$, and

$$\min(1, k(f-a)^+) \uparrow_k I_Z.$$

By Lemma 3.1.7.(v), $I_Z \in L'$, hence $Z \in \mathscr{C}$. □

Lemma 3.1.10. *The σ-fields $\sigma(L), \sigma(L')$ and $\sigma(\mathscr{C})$ are identical.*

Proof. By Lemma 3.1.9, every function in L' is measurable. Therefore, $\sigma(L') \subseteq \sigma(\mathscr{C})$. The other inclusion is immediate by the definition of $\mathscr{C}$: if $G \in \mathscr{C}$, then $I_G \in L'$; hence $G = \{I_G = 1\} \in \sigma(L')$. Thus, $\mathscr{C} \subseteq \sigma(L')$, from where $\sigma(\mathscr{C}) \subseteq \sigma(L')$.
Now let $f \in L'$ and consider a sequence in L with $f_n \uparrow f$. Since every f_n is $\sigma(L)$-Borel measurable, it follows that their pointwise limit f is $\sigma(L)$-Borel measurable. Thus, $\sigma(L') \subseteq \sigma(L)$, because $\sigma(L')$ is the smallest σ-field making every $f \in L'$ Borel measurable.
If $f \in L$, we can split $f = f^+ - f^-$. Note that, since L is closed under the lattice operations and contains all constant functions, then both f^+ and f^- are in L. Now $f^+, f^- \in L^+ \subseteq L'$, and thus f is $\sigma(L')$-Borel measurable. It follows that $\sigma(L) \subseteq \sigma(L')$. □

Lemma 3.1.11. *For any $A \subseteq \Omega$, $\mu^*(A) = \inf\{E(f) \,|\, f \in L', f \geq I_A\}$.*

Proof. By definition of μ, $\mu^*(A) = \inf\{E(I_G) | G \in \mathscr{C}, A \subseteq G\}$. It is then clear that $\mu^*(A) = \inf\{E(f) \,|\, f = I_G, G \in \mathscr{C}, f \geq I_A\} \geq \inf\{E(f) \,|\, f \in L', f \geq I_A\}$. To see the other inequality, let f be a function in L' with $f \geq I_A$, and a real number in the interval $a \in (0,1)$. Let $Z = \{f > a\}$. Note that, since f is measurable, we have $Z \in \sigma(\mathscr{C})$. Additionally, $f \geq I_A$ and since I_A is 1 on A, it follows that $A \subseteq Z$. Thus $\mu^*(A) \leq \mu(Z) = E(I_Z)$. Now note that $f \geq a I_Z$, and therefore $E(f) \geq a E(I_Z)$. Finally, we have $\mu^*(A) \leq \frac{E(f)}{a}$. The result is deduced by letting $a \to 1^-$. □

Lemma 3.1.12. *If $\mathscr{H} = \{H \subseteq \Omega \,|\, \mu^*(H) + \mu^*(H^c) = 1\}$, then $\mathscr{G} \subseteq \mathscr{H}$. Therefore, $\sigma(\mathscr{C}) \subseteq \mathscr{H}$.*

Proof. Let $G \in \mathscr{C}$. Since $I_G \in L'$ by definition, we can find a sequence of functions in L^+ such that $f_n \uparrow I_G$. On one side, this implies that $E(f_n) \uparrow E(I_G) = \mu(G) = \mu^*(G)$. On the other, it imples that $1 - f_n \downarrow 1 - I_G = I_{G^c} \geq 0$, and then $1 - f_n \in L^+ \subseteq L'$. Hence, by Lemma 3.1.11,

$$\mu^*(G^c) = \inf\{E(f) \,|\, f \in L', f \leq I_{G^c}\} \geq \inf_n E(1 - f_n) = 1 - \lim_n E(f_n) = 1 - E(I_G).$$

Therefore, $\mu^*(G) + \mu^*(G^c) \leq 1$. Since the inequality $\mu^*(G) + \mu^*(G^c) \geq 1$ always holds by Lemma 2.3.3.(iii), we have $G \in \mathscr{H}$. □

Finally, we are in position to obtain the main result of the section, and a useful corollary. For clarity, the hypotheses accumulated so far are gathered in the statement of the theorem.

Theorem 3.1.13 (Daniell Representation Theorem). *Let L be a vector space of real-valued functions on the set Ω that contains all constant functions and is closed under lattice operations. Let E be a Daniell integral on L such that $E(1) = 1$.*
Then, there is a unique probability measure P on $\sigma(L)$ such that each $f \in L$ is P-integrable and

$$E(f) = \int_\Omega f \, dP.$$

Proof. Let P be the restriction of μ^* to $\sigma(L)$. By Lemma 3.1.8 and Lemma 3.1.12, P is a probability measure on $\sigma(\mathscr{C})$, which is equal to $\sigma(L)$ by Lemma 3.1.10.
As is usual when working with measures, we will show the result for indicators first and then work upwards. However, a specific detail needs to be addressed: we cannot consider any indicator function I_G, with $G \in \sigma(\mathscr{C})$, since if $G \in \sigma(\mathscr{C}) \setminus \mathscr{C}$, then $I_G \notin L'$, and then $E(I_G)$ need not be defined. However, the result clearly holds for indicators of sets $B \in \mathscr{C}$:

$$E(I_B) = \mu^*(B) = P(B) = \int_\Omega I_B \, dP.$$

By additivity, it also holds for (real) linear combinations of such indicators (a subclass of simple functions). However, this class of functions is enough to approximate - via limits - all functions in L^+: take some $f \in L^+$, and consider the sequence of functions

$$s_n = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} I_{\{(i-1)/2^n < f \leq i/2^n\}} + n I_{\{f>n\}}.$$

Note that since $I_{\{(i-1)/2^n < f \leq i/2^n\}} = I_{\{f>(i-1)/2^n\}} - I_{\{f>i/2^n\}}$ and $\{f > a\} \in \mathscr{C}$ (Lemma 3.1.9), we can consider $E(s_n)$ and $E(s_n) = \int_\Omega s_n \, dP$. Additionally, since $s_n \uparrow f$, following the Monotone Convergence Theorem and the fact that E is a Daniell integral, we have $E(f) = \lim_n E(s_n) = \lim_n \int_\Omega s_n \, dP = \int_\Omega f \, dP$.
If $f \in L$, split $f = f^+ - f^-$. Since L is closed under the lattice operations and contains all constant functions (the function 0, in particular), we have $f^+, f^- \in L$. Then, by additivity,

$$E(f) = E(f^+) - E(f^-) = \int_\Omega f^+ \, dP - \int_\Omega f^- \, dP = \int_\Omega f \, dP.$$

(Since E is real-valued, both integrals are finite).
To see uniqueness, consider another such probability measure P' and note that the class of sets of $\sigma(\mathscr{C})$ where they coincide is a D-system and contains $\mathscr{C}$, because $\int_\Omega f \, dP = \int_\Omega f \, dP'$ for all $f \in L$; hence on L'. By the definition of $\mathscr{C}$, $I_G \in L'$ for all $G \in \mathscr{C}$; thus, $P(G) = \int_\Omega I_G \, dP = \int_\Omega I_G \, dP' = P'(G)$. Finally, by Corollary 3.1.3, P and P' agree on all of $\sigma(\mathscr{C})$. $\square$

Finally, we give an approximation theorem.

Theorem 3.1.14. *Under the hypotheses and notation of Theorem 3.1.13, assume in addition that L is closed under limits of uniformly convergent sequences of functions. Let*

$$\mathscr{C}' = \left\{ G \subseteq \Omega \,\middle|\, G = \{f(\omega) > 0\} \text{ for some } f \in L^+ \right\}.$$

Then,

(i) $\mathscr{C}' = \mathscr{C}$

(ii) *If* $A \in \sigma(L)$*, then* $P(A) = \inf\{P(G) \,\big|\, G \in \mathscr{C}', A \subseteq G\}$.

(iii) *If* $G \in \mathscr{C}$*, then* $P(G) = \sup\{E(f) \,\big|\, f \in L^+, f \leq I_G\}$

Proof. 1. First note that $\mathscr{C}' \subseteq \mathscr{C}$ by Lemma 3.1.9. Conversely, take $G \in \mathscr{C}$ and consider a sequence in L^+ with $f_n \uparrow I_G$. Define $f = \sum_n 2^{-n} f_n$. Since $0 \leq f_n \leq 1$, the series is uniformly convergent, hence $f \in L^+$. Now, $f(\omega) = 0$ if, and only if, $f_n(\omega) = 0$ for all n. Thus,

$$\{f(\omega) > 0\} = \bigcup_n \{f_n(\omega) > 0\} = \{I_G(\omega) > 0\} = G. \qquad \square$$

Consequently, $G \in \mathscr{C}$.

2. Immediate from 3.1.14.(i) and the fact that $P = \mu^*$ on $\sigma(L)$ (using the definition of μ^*).

3. If $f \in L^+$ and $f \leq I_G$, then $E(f) \leq E(I_G) = P(G)$. Conversely, let $G \in \mathscr{C}$ and consider a sequence in L^+ with $f_n \uparrow I_G$. Then $f_n \leq f$ and $P(G) = E(I_G) = \lim_n E(f_n) = \sup_n E(f_n)$, hence $P(G) \leq \sup\{E(f) \,\big|\, f \in L^+, f \leq I_G\}$.

3.2 Measure and Topology

In this section, we will study how the Measure Theory developed so far relates to Topology. The main goal is to obtain a result which allows us to approximate probabilities of Borel sets by probabilities of compact sets. This will be a key tool in our proof of the Kolmogorov Extension Theorem, which is the main goal of this work.

Definition 3.2.1. *Let* Ω *be a* ***normal*** *topological space, that is,* Ω *is Hausdorff and for every two disjoint closed subsets* $A, B \subseteq \Omega$*, there exist two disjoint open sets* $U, V \subseteq \Omega$ *such that* $A \subseteq U$ *and* $B \subseteq V$.
We will denote the class of all continuous functions from Ω *to* $\mathbb{R}$ *(with the standard topology) as* $C(\Omega)$*, and the class of all such functions that are, additionally, bounded, as* $C_b(\Omega)$.

The basic property we will need about normal topological spaces is *Urysohn's Lemma*. It has an admittedly technical proof, but it uses only the definition and elementary topology notions.

Theorem 3.2.2 (Urysohn's Lemma). *If* A *and* B *are disjoint closed subsets of a normal topological space* Ω*, there exists a continuous function* $f : \Omega \to [0,1]$ *such that* $f = 0$ *on* A *and* $f = 1$ *on* B.

Proof. In this proof, D' and D will denote the sets of *Dyadic rationals* in $(0,1)$ and $[0,1]$, respectively; a Dyadic rational is a rational number of the form $\frac{k}{2^n}$, with $n, k \in \mathbb{Z}$.
An immediate characterisation of normality is the following: if U is an open set, V is closed and $V \subseteq U$, then there exist an open set U' and a closed set V' such that $V \subseteq U' \subseteq V' \subseteq U$.

We will use this characterisation to show that there exists a family of open sets $U(r)$ and closed sets $V(r)$, where $r \in D'$, such that

- $A \subseteq U(r) \subseteq V(r) \subseteq B^c$ for any $r \in D'$,
- If $r, s \in D'$ with $r < s$, then $V(r) \subseteq U(s)$.

Extend this notation from D' to D as $U(1) = B^c$, $V(0) = A$. Note that any $r \in D$ can be written as $\frac{k}{2^n}$, where n is a positive integer and $k = 0, \ldots, 2^n$. This allows us to proceed our construction by induction on n:

For $n = 1$, note that B^c is open and $A \subseteq B^c$. Hence, applying the characterisation of normality, there exist an open set $U\left(\frac{1}{2}\right)$ and a closed set $V\left(\frac{1}{2}\right)$ such that $A \subseteq U\left(\frac{1}{2}\right) \subseteq V\left(\frac{1}{2}\right) \subseteq B^c$.

For the inductive step, consider some integer $0 \leq k \leq 2^{n+1}$. If k is even, we can write $\frac{k}{2^{n+1}} = \frac{k'}{2^n}$ (where $k = 2k'$), and in this case the construction is already made. If k is odd, write it as $k = 2k' + 1$. Note that $0 \leq k' < 2^n$. Then, by the construction made so far yields us the sets $V\left(\frac{k'}{2^n}\right)$ and $U\left(\frac{k'+1}{2^n}\right)$, which satisfy $V\left(\frac{k'}{2^n}\right) \subseteq U\left(\frac{k'+1}{2^n}\right)$. Apply the characerisation of normality to these sets (the former is closed and the latter is open) to obtain two intermediate sets U' (open) and V' (closed). Define $U\left(\frac{k}{2^{n+1}}\right) = U'$ and $V\left(\frac{k}{2^{n+1}}\right) = V'$.

Having constructed the last class of sets, define

$$f(\omega) = \begin{cases} 1, & \text{if } \omega \not\in U(1), \\ \inf\{r \in D' \,|\, \omega \in U(r)\}, & \text{otherwise.} \end{cases}$$

It is clear that $f = 0$ on A, $f = 1$ on B and that $0 \leq f \leq 1$. We need only to show that f is continuous. Note that, for any given $r \in D'$, $\omega \in U(r)$ implies $f(\omega) \leq r$ (this is immediate); and $\omega \in V(r)^c$ implies $f(\omega) \geq r$. This last statement can be shown by contradiction: suppose that $f(\omega) < r$. It is impossible that $f(\omega) = 1$, because $r \leq 1$. Therefore, $f(\omega) = \inf\{s \in D' \,|\, \omega \in U(s)\}$, hence there exists some $s \in D'$ such that $s < r$ and $\omega \in U(s)$. But $U(s) \subseteq V(s) \subseteq U(r) \subseteq V(r)$, so that $\omega \in V(r)$, a contradiction. To see continuitiy, take some $\omega \in \Omega$ and $\varepsilon > 0$.

If $f(\omega) = 0$, take some $r \in D'$ such that $r < \varepsilon$ and $\omega \in U(r)$. Then, $U(r)$ is an open neighbourhood of ω such that $f(U(r)) \subseteq [0, \varepsilon)$. If $f(\omega) = 1$, choose some $r \in D$ such that $1 - \varepsilon < r$. Then, $V(r)^c$ is an open neighbourhood of ω such that $f(V(r)^c) \subseteq (1 - \varepsilon, 1]$. If $0 < f(\omega) < 1$, choose some $r, s \in D'$ so that $f(\omega) - \varepsilon < r < f(\omega) < s < f(\omega) + \varepsilon$. Then, $U(s) \setminus V(r)$ is an open neighbourhood of ω such that $f(U(s) \setminus V(r)) \subseteq (f(\omega) - \varepsilon, f(\omega) + \varepsilon)$[3]. □

Definition 3.2.3. *The class of* Baire sets *of* Ω*, denoted by* $\mathscr{A}(\Omega)$*, is the smallest* σ*-field on* Ω *making all continuous real-valued functions Borel measurable. Namely,*

$$\mathscr{A}(\Omega) = \sigma(C(\Omega)) = \sigma\left(\left\{f^{-1}(B) \,\middle|\, f \in C(\Omega) \text{ and } B \text{ is open in } \mathbb{R}\right\}\right)$$

A ***σ-closed*** *set is a countable union of closed sets (which need not be closed nor open) and a* ***σ-open*** *set is a countable intersecion of open sets (which need not be closed nor open either).*

[3]A proof for this result was not included in [6]. The version presented here is somewhat original: various ideas were taken from online forums.

Remark 3.2.4. *It is immediate from the definition that* $\mathscr{A}(\Omega) \subseteq \mathscr{B}(\Omega)$*: every* $f^{-1}(B)$*, with* f *continuous and* B *open, is open; hence Borel. Then, apply Remark 2.2.2.*

We can obtain a smaller class of generators of the Baire sets by considering only bounded functions (we are claiming that $\sigma(C_b(\Omega)) = \sigma(C(\Omega)) = \mathscr{A}(\Omega)$*): suppose that all bounded, continuous functions are measurable. If* $f \in C(\Omega)$*, then for each* $n \in \mathbb{N}$*, the function* $\max(f, n)$ *is continuous: it is the composition of* f *and the continuous mapping* $x \mapsto \max(x, n)$*. We can apply the result for* $n = 0$ *to obtain that* $f^+ = \max(f, 0)$ *is continuous. Then,* $\max(f^+, n) \in C_b(\Omega)$ *is also continuous, hence measurable. Since* $\max(f^+, n) \uparrow f^+$*, it follows that* f^+ *is measurable. A similar argument used with* $f^- = -\max(-f, 0)$ *yields that* f^- *is measurable. Finally,* $f = f^+ - f^-$ *is measurable.*

Lemma 3.2.5. *Let* Ω *be a normal topological space. Then* $\mathscr{A}(\Omega)$ *is the smallest* σ*-field containing the* σ*-open sets that are also closed (or equally well, the* σ*-closed sets that are also open).*

Proof. Let $\mathscr{H}$ be the σ-field generated by all open, σ-closed sets. Let $f \in C(\Omega)$. Then,

$$\{f > a\} = \bigcup_n \left\{ f \geq a + \frac{1}{n} \right\},$$

hence $\{f > a\}$ is an open, σ-closed set. It follows that f is $\mathscr{H}$-measurable. Since $\mathscr{A}(\Omega)$ is the smallest σ-field making all such functions measurable (by Remark 3.2.4), it follows that $\mathscr{A}(\Omega) \subseteq \mathscr{H}$. To see the other inclusion, let $H = \bigcup_n F_n$ be an open, σ-closed set (F_n is a closed set for each n). Note that, for each n, H^c and F_n are disjoint closed sets. Use Urysohn's Lemma to obtain a function f_n such that $f_n = 0$ on H^c and $f_n = 1$ on F_n. Then, define the function $f = \sum_n 2^{-n} f_n$. This series is uniformly convergent (the Weierstrass M test can be used to see this) and bounded by 1 since $0 \leq f_n \leq 1$, hence $f \in C_b(\Omega)$ and $f \geq 0$. Additionally,

$$\{f > 0\} = \bigcup_n \{f_n > 0\} = H.$$

Therefore, $H \in \mathscr{A}(\Omega)$, so that $\mathscr{H} \subseteq \mathscr{A}(\Omega)$. □

We can examine the last argument to obtain another result:

Corollary 3.2.6. *If* Ω *is a normal topological space, the open,* σ*-closed sets are precisely the sets* $\{f > 0\}$ *with* $f \in C_b(\Omega)$*,* $f \geq 0$*.*

Proof. In the first part of the proof, consider the special case $f \in C_b(\Omega)$ and $a = 0$. Then, $\{f > 0\}$ is shown there to be an open, σ-closed set.

In the reciprocal part the proof, we showed that if H is an open, σ-closed set, then H can be written as $\{f > 0\}$, where $f \in C_b(\Omega)$ and $f \geq 0$. □

Lemma 3.2.7. *Let* A *be an open,* σ*-closed set in the normal space* Ω*. Then,* I_A *is the limit of an increasing sequence of continuous functions.*

Proof. Let $f = I_A$. Use Corollary 3.2.6 to write $A = \{f > 0\}$, and define $A_n = \{f \geq \frac{1}{n}\}$, so that $A_n \uparrow A$. Use Urysohn's Lemma to obtain a sequence of functions $0 \leq f_n \leq 1$ such that $f_n = 0$ on A^c and $f_n = 1$ on A_n. Define $g_n = \max(f_1, \ldots, f_n)$, so that $0 \leq g_n \leq 1$. The functions g_n are clearly continuous and form an increasing sequence.

Furthermore, they satisfy $I_{A_n} \leq g_n \leq I_A$: outside of A, all f_n are 0, and thus so is g_n. Inside of A_n, $f_n = 1$, hence $g_n = 1$. Taking limits, since $I_{A_n} \uparrow I_A$, it follows that $g_n \uparrow I_A$ too. □

At this point, we are close to obtaining our approximation theorem, but first we need to be able to approximate by closed sets. The Daniell Theory provides useful tools for this purpose.

Theorem 3.2.8. *Let P be any probability measure on $\mathscr{A}(\Omega)$, where Ω is a normal topological space. If $A \in \mathscr{A}(\Omega)$, we have*

(i) $P(A) = \inf\{P(V) \mid A \subseteq V$ *and V is an open, σ-closed set*$\}$.

(ii) $P(A) = \sup\{P(C) \mid C \subseteq A$ *and C is a closed, σ-open set*$\}$.

Proof. Define $L = C_b(\Omega)$, so that $\mathscr{A}(\Omega) = \sigma(L)$, and $E(f) = \int_\Omega f\, dP$. It is clear that L is closed under the lattice operations and that $E(f) = \int_\Omega f\, dP$ exists and is finite because f is bounded and P is a finite measure. By the Monotone Convergence Theorem, E is a Daniell integral: if $f_n \uparrow f$, $f_n \geq 0$, then $E(f_n) \uparrow E(f)$; and if $f_n \downarrow 0$, then $E(f_n) \downarrow E(0) = 0$.

Therefore, we can use Theorem 3.1.14.(ii) to see that

$$P(A) = \inf\{P(G) \mid A \subseteq G, G \in \mathscr{C}'\},$$

where $\mathscr{C}' = \{G \subseteq \Omega \mid G = \{f > 0\}, f \in L^+\}$. By Corollary 3.2.6, $\mathscr{C}'$ is exactly the class of all open, σ-closed sets.

The proof for 3.2.8.(ii) is obtaned by applying the result seen so far to A^c and taking into account that $P(B^c) = 1 - P(B)$ for every $B \in \mathscr{A}(\Omega)$. □

In metric spaces (which are always normal), we can obtain stronger results:

Proposition 3.2.9. *If Ω is a metric space, then every closed set is a σ-open set. Therefore, $\mathscr{A}(\Omega) = \mathscr{B}(\Omega)$.*

Proof. Consider a closed subset $F \subseteq \Omega$. Define the *distance to F* function $\rho(\omega) = \operatorname{dist}(\omega, F) = \inf_{f \in F} d(\omega, f)$. Then, ρ is continuous:

Let f be an arbitrary point in F. Let $\omega_1, \omega_2 \in \Omega$. Call $\delta = d(\omega_1, \omega_2)$. Then,

$$\rho(\omega_2) \leq d(\omega_2, f) \leq d(\omega_1, f) + \delta.$$

Since f is arbitrary, then $\rho(\omega_2) \leq \rho(\omega_1) + \delta$. Switching papers for ω_1 and ω_2, one obtains that $\rho(\omega_1) \leq \rho(\omega_2) + \delta$. Therefore, if we take $\delta = \varepsilon$, we conclude that ρ is continuous as mapping between metric spaces.

From this, it follows that F is σ-open set:

$$F = \{\omega \in \Omega \mid \rho(\omega) = 0\} = \bigcap_n \left\{\omega \in \Omega \,\middle|\, \rho(\omega) < \frac{1}{n}\right\}.$$

(in the first equality, we used that F is closed). The inclusion $\mathscr{B}(\Omega) \subseteq \mathscr{A}(\Omega)$ now follows from Lemma 3.2.5. The reciprocal inclusion was already established in Remark 3.2.4.□

From this last theorem we can, of course, deduce that in a metric space every open set is a σ-closed set. Combining these facts with Theorem 3.2.8, we obtain a simpler statement:

Corollary 3.2.10. *Let P be any probability measure on $\mathscr{A}(\Omega)$, where Ω is a metric space. If $A \in \mathscr{A}(\Omega)$, we have*

(i) $P(A) = \inf\{P(V) \,|\, A \subseteq V$ *and* V *is an open set*$\}$.

(ii) $P(A) = \sup\{P(C) \,|\, C \subseteq A$ *and* C *is a closed set*$\}$.

Finally, we can state the desired result:

Theorem 3.2.11 (Approximation by compact sets). *Let Ω be a complete, separable metric space. If P is a probability measure on $\mathscr{B}(\Omega)$, then, for each $A \in \mathscr{B}(\Omega)$,*

$$P(A) = \sup\left\{P(K) \,\middle|\, K \text{ is a compact subset of } A\right\}.$$

Proof. We will first show that, for every $\varepsilon > 0$, there exists some compact set K_ε such that $P(K_\varepsilon) > 1 - \frac{\varepsilon}{2}$.

Since Ω is separable, there exists a sequence of points ω_m that is dense in Ω. Consider some arbitrary radius $r > 0$, the open balls $B_m(r) = B(\omega_m, r)$ and their closures $\overline{B}(\omega_m, r)$. Then, $\Omega = \bigcup_m \overline{B}(\omega_m, r)$ for any given radius r. Write $U_{nm} = \bigcup_{k=1}^{m} \overline{B}\left(\omega_k, \frac{1}{n}\right)$, so that, for every $n \in \mathbb{Z}^+$, we have $U_{nm} \uparrow_m \Omega$. It follows that there exists some $m(n) \in \mathbb{Z}^+$ such that $P\left(U_{nm(n)}\right) \geq P(\Omega) - \varepsilon 2^{-n-1} = 1 - \varepsilon 2^{-n-1}$. Now define $K_\varepsilon = \bigcap_n U_{nm(n)}$.

Firstly, note that

$$P(K_\varepsilon^c) = P\left(\bigcup_n U_{nm(n)}^c\right) \leq \sum_n P\left(U_{nm(n)}^c\right) = \sum_n 1 - P\left(U_{nm(n)}\right) \leq \sum_n \varepsilon 2^{-n-1} = \frac{\varepsilon}{2}.$$

It remains to show that K_ε is compact. In a metric space, compactness and sequential compactness are equivalent. Therefore, it suffices to show that every sequence in K_ε has a subsequence converging to a point in K_ε. Since K_ε is clearly closed (it is the intersection of closed sets), if we show that any sequence in K_ε has a converging subsequence, its limit will automatically be in K_ε. One last simplification can be made: since Ω is complete, it suffices to show that every sequence in K_ε has a Cauchy subsequence.

Let $\{x_p\}_{p \in \mathbb{Z}^+}$ be a sequence of points in K_ε. First, note that, since $x_p \in U_{1m(1)} = \bigcup_{k=1}^{m(1)} \overline{B}(\omega_k, 1)$, there exists some $k_1 \leq m(1)$ such that infinitely many x_p are in $\overline{B}(\omega_{k_1}, 1)$. Let T_1 be the (infinite) set of all such indices. Similarly, $x_p \in \bigcup_{k=1}^{m(2)} \overline{B}\left(\omega_k, \frac{1}{2}\right)$ for all $p \in T_1$; hence there exists some $k_2 \leq m(2)$ such that infinitely many values x_p, with $p \in T_1$, are in $\overline{B}\left(\omega_{k_2}, \frac{1}{2}\right)$. Let T_2 be the (infinite) set of all such indices. Continue inductively to obtain integers $k_1, k_2, \ldots$ and infinite sets of indices $T_1 \supseteq T_2 \supseteq \ldots$ so that, for every i,

$$x_p \in \bigcap_{j=1}^{i} \overline{B}\left(\omega_{k_j}, \frac{1}{j}\right) \text{ for all } p \in T_i$$

Choose one $p_i \in T_i$ in a manner such that $p_1 < p_2 < \ldots$ (this is always possible because the sets T_i are infinite), and consider the subsequence $x_{p_1}, x_{p_2}, \ldots$. Then, if $l > j$, both x_{p_j} and x_{p_l} are in $\overline{B}\left(\omega_{k_j}, \frac{1}{j}\right)$; thus,

$$d(x_{p_j}, x_{p_l}) \leq \frac{2}{j} \to 0 \text{ when } j \to +\infty.$$

We have now obtained the desired compact set K_ε. To see the result we wanted, note that any compact subset K of A satisfies $P(K) \le P(A)$. It follows that

$$P(A) \ge \sup\left\{P(K) \,\middle|\, K \text{ is a compact subset of } A\right\}.$$

To see the equality, note that by 3.2.10.(ii), the result holds when "compact" is replaced with "closed". Therefore, for every $\varepsilon > 0$, there exists some closed set $C \subseteq A$ such that $P(A) - P(C) \le \frac{\varepsilon}{2}$. Take $K = C \cap K_\varepsilon$.
Note that $P(C) - P(K) = P\left(C \setminus (C \cap K_\varepsilon)\right) = P(C \setminus K_\varepsilon)) \le P(K_\varepsilon^c) < \frac{\varepsilon}{2}$. Therefore, it follows that K is a compact subset of A and

$$P(A) - P(K) = P(A) - P(C) + P(C) - P(K) \le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

finishing the proof. □

To end the section, we offer a short lemma regarding separable metric spaces.

Lemma 3.2.12. *In metric spaces, separability and second countability are equivalent: that is, if Ω is a metric space, then there exists a countable, dense subset $S \subseteq X$ if, and only if, there exists a countable basis for X*[4].

Proof. First suppose that X is separable, and write $S = \{x_n\}_{n \in \mathbb{Z}^+}$. Define, for each $n, m \in \mathbb{Z}^+$, the open set

$$B_{nm} = B\left(x_n, \frac{1}{m}\right).$$

We will show that the set $\{B_{nm}\}_{n,m \in \mathbb{Z}^+}$ forms a basis for the topology in X. Let U be an open set, and define

$$U' = \bigcup_{a,b} B_{ab},$$

where a and b range over the pairs of positive integers (n, m) such that $B_{nm} \subseteq U$. It is then clear that $U' \subseteq U$. To see the other inclusion, take some $z \in U$. Then, there exists some $\varepsilon > 0$ such that $B(z, \varepsilon) \subseteq U$. By the Archimedian Property, there exists some $m \in \mathbb{Z}^+$ such that $\frac{1}{m} < \frac{\varepsilon}{2}$. Since S is dense in X, there exists some x_n such that $d(x_n, z) < \frac{1}{m}$. Finally, $z \in B_{nm} \subseteq U$, and thus $z \in U'$.
For the reciprocal implication, consider some countable basis $\{B_n\}_{n \in \mathbb{Z}^+}$ and choose some $x_n \in B_n$ for each n. Then, the set $S = \{x_n\}_{n \in \mathbb{Z}^+}$ is trivially dense in X. □

[4] This result is original.

CHAPTER 4

PRODUCT SPACES

Often times in topology or algebra, when working with a structure (be it a topological space, a group, a ring...) it is interesting or convenient to be able to create *product* structures; that is, equip a cartesian product with the same structure their components had[1]. In this chapter, we aim to construct said product in measure spaces.

It is convenient to warn the reader that the notation used in this chapter is, to the writer's knowledge, original, and in any case different to that found in [6].

4.1 Finite product of measure spaces

We will begin our study of product measure spaces with the finite case. As we will see, the *algebraic* part of this - the part regarding measurability and σ-fields - is rather easy and done very similarly to how it is done in topology. In the *analytical* details - that is, when measures take play - is where most of the theory arises.

We will develop first the *algebraic* part: the structures where we will work. As in topology, it is desirable that cartesian products of sets in the base structures also belong to the product structure.

Definition 4.1.1. *Let $\mathscr{F}_j$ be a σ-field of subsets of some set Ω_j, $j = 1, \ldots, n$. Define $\Omega = \Omega_1 \times \ldots \times \Omega_n$. A **measurable rectangle** of $\mathscr{F}_1, \ldots, \mathscr{F}_n$, or, for short, a **rectangle** in Ω is a set $A \subseteq \Omega$ of the form $A = A_1 \times \ldots \times A_n$, with $A_i \in \mathscr{F}_i$.*

*The smallest σ-field over Ω containing all measurable rectangles of $\mathscr{F}_1, \ldots, \mathscr{F}_n$ is called the **product σ-field**, and is denoted by $\mathscr{F}_1 \times \ldots \times \mathscr{F}_n$ (this is not the cartesian product of the sets $\mathscr{F}_1, \ldots, \mathscr{F}_n$). If $\mathscr{F}_1 = \ldots = \mathscr{F}_n = \mathscr{F}$ for some σ-field $\mathscr{F}$, we denote $\mathscr{F}^n = \mathscr{F}_1 \times \cdots \times \mathscr{F}_n$.*

Consider a finite, proper subset $v \subseteq \{1, \ldots, n\}$. Write $v = \{i_1, \ldots, i_r\}$, and $v^c = \{j_1, \ldots, j_l\}$ (then, $n = r + l$), with $i_1 < \ldots < i_r$ and $j_1 < \ldots < j_l$. Fix the values $\omega_{i_1}, \ldots, \omega_{i_r}$, with

[1] The categorical process that justifies this is the *product universal property*.

*$\omega_{i_k} \in \Omega_{i_k}$. Define the **section** of a given set $F \subseteq \Omega$ at $\omega_{i_1}, \ldots, \omega_{i_r}$ as*

$$F^{\omega_{i_1},\ldots,\omega_{i_r}} = \left\{ \left(\omega_{j_1}, \ldots, \omega_{j_l}\right) \in \Omega_{j_1} \times \cdots \times \Omega_{j_l} \,|\, (\omega_1, \ldots, \omega_n) \in F \right\}.$$

*Finally, define the **section** of a given function $f : \Omega \to X$ (where X is some arbitrary set) at $\omega_{i_1}, \ldots, \omega_{i_r}$ as the function $f^{\omega_{i_1},\ldots,\omega_{i_r}} : \Omega_{j_1} \times \cdots \times \Omega_{j_l} \to X$ Such that $(\omega_{j_1}, \ldots, \omega_{j_l}) \mapsto f(\omega_1, \ldots, \omega_n)$.*

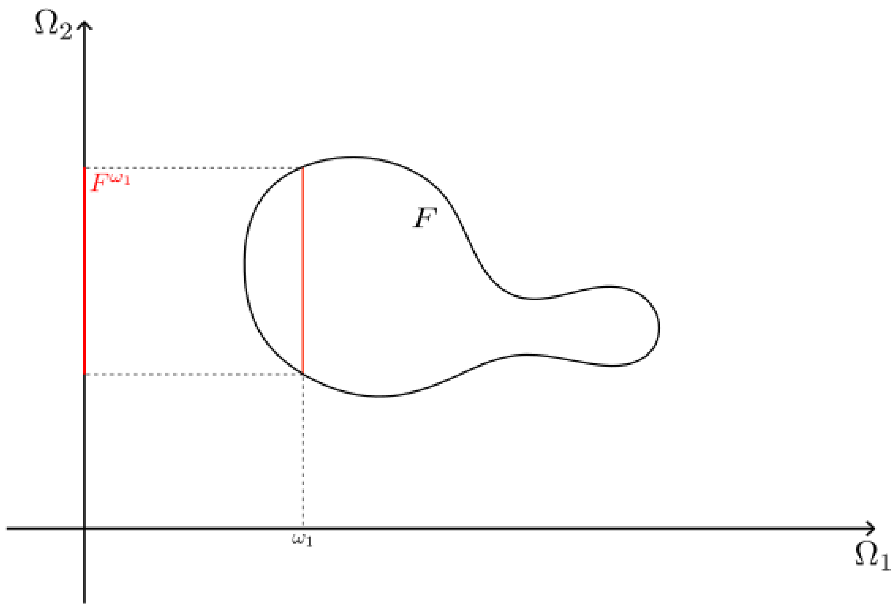

Figure 4.1: Section F^{ω_1} of a subset F of a product of two measurable spaces, Ω_1 and Ω_2.

Sections have very interesting properties regarding set operations and measurable spaces.

Proposition 4.1.2 (Properties of sections). *Under the notation of the previous definition, let $\{A_t \,|\, t \in T\}$ be an arbitrarily indexed collection of subsets of $\Omega = \Omega_1 \times \cdots \times \Omega_n$.*[2] *Then,*

(i) For every fixed values $\omega_{i_1}, \ldots, \omega_{i_r}$, with $\omega_{i_k} \in \Omega_{i_k}$, one has

$$\left(\bigcup_{t\in T} A_t\right)^{\omega_{i_1},\ldots,\omega_{i_r}} = \bigcup_{t\in T} A_t^{\omega_{i_1},\ldots,\omega_{i_r}} \quad \textit{and} \quad \left(\bigcap_{t\in T} A_t\right)^{\omega_{i_1},\ldots,\omega_{i_r}} = \bigcap_{t\in T} A_t^{\omega_{i_1},\ldots,\omega_{i_r}}.$$

Additionally, for every set $A \in \Omega$,

$$\left(A^{\omega_{i_1},\ldots,\omega_{i_r}}\right)^c = \left(A^c\right)^{\omega_{i_1},\ldots,\omega_{i_r}}.$$

(ii) For every fixed values $\omega_{i_1}, \ldots, \omega_{i_r}$, with $\omega_{i_k} \in \Omega_{i_k}$, $\emptyset^{\omega_{i_1},\ldots,\omega_{i_r}} = \emptyset$. In particular, if the sets A_t are disjoint pairwise, so are the sections $A_t^{\omega_{i_1},\ldots,\omega_{i_r}}$.

[2]This result is original.

(iii) Sections of measurable sets are always measurable in their respective σ-fields.

(iv) Sections of measurable functions are always measurable in their respective σ-fields.

Proof. 1. This proof is a simple verification of quantifiers. We will omit it.

2. Clearly, a section of the empty set is empty. Moreover, if A and B are disjoint subsets of Ω, then, by 4.1.2.(i),

$$A^{\omega_{i_1},\dots,\omega_{i_r}} \cap B^{\omega_{i_1},\dots,\omega_{i_r}} = (A \cap B)^{\omega_{i_1},\dots,\omega_{i_r}} = \emptyset^{\omega_{i_1},\dots,\omega_{i_r}} = \emptyset.$$

3. Write $\mathscr{F} = \mathscr{F}_1 \times \ldots \times \mathscr{F}_n$, and $\mathscr{F}_v = \mathscr{F}_{j_1} \times \ldots \times \mathscr{F}_{j_l}$.

   Define the class of sets $\mathscr{C} = \{B \in \mathscr{F} | F^{w_{i_1},\dots,w_{i_r}} \in \mathscr{F}_v\}$. It is clear that $\mathscr{C}$ is a σ-field, because of 4.1.2.(i) and the fact that $\mathscr{F}_v$ is a σ-field. Additionally, $\mathscr{C}$ contains all measurable rectangles, because

$$(A_1 \times \ldots \times A_n)^{w_{i_1},\dots,w_{i_r}} = \begin{cases} A_{j_1} \times \ldots \times A_{j_l}, & \text{if } (w_{i_1},\dots,w_{i_r}) \in A_{i_1} \times \ldots \times A_{i_r}, \\ \emptyset, & \text{otherwise.} \end{cases}$$

   Thus, $\mathscr{F} = \mathscr{C}$.

4. This is an immediate consequence of 4.1.2.(iii) and the fact that, for any given set $Z \subseteq X$,

$$\left(f^{\omega_{i_1},\dots,\omega_{i_r}}\right)^{-1}(Z) = \left(f^{-1}(Z)\right)^{\omega_{i_1},\dots,\omega_{i_r}}. \qquad \square$$

Remark 4.1.3. *In a given product measurable space, the class of disjoint unions of measurable rectangles forms a field.*

We have now covered the *algebraic* part of the section. We have developed the structure on which we will measure, and now wish to construct measures on these product spaces using measures on the base spaces.

The following theorem does this for the $n = 2$ case. The intuition behind it can be explained in terms of Figure 4.1. If we want to measure a set $F \subseteq \Omega_1 \times \Omega_2$ we can consider, for every $\omega_1 \in \Omega_1$, the intersection between the vertical line at ω_1 and the set F. This intersection is parallel to the section F^{ω_1}, which we are able to measure since it is a measurable subset of Ω_2. The way we measure F^{ω_1} can be the same for each ω_1[3], or can vary depending on it, so that we have a different measure μ^{ω_1} on Ω_2 for every ω_1. Finally, to measure F we can integrate all of the different values obtained for each ω_1 with respect to a given measure on Ω_1. Finally, we need to impose some technical constraints regarding measurability and σ-finiteness, which are included in the statement of the result.

Theorem 4.1.4 (Two-Dimensional Product Measure Theorem). *Let* $(\Omega_1, \mathscr{F}_1)$ *and* $(\Omega_2, \mathscr{F}_2)$ *be measurable spaces, and μ a σ-finite measure on $\mathscr{F}_1$. Assume that we are given a map $\mu \colon \Omega_1 \times \mathscr{F}_2 \to \overline{\mathbb{R}}$ which satisfies:*

- *The mapping $\mu^{\omega_1} = \mu(\omega_1, \cdot)$ is a measure over $\mathscr{F}_2$ for every fixed $\omega_1 \in \Omega_1$.*

[3]This is the approach followed in the "classical" Product Measure Theorem; see Corollary 4.1.8.

- *$\mu(\cdot, B)$ is $\mathscr{F}_1$-measurable for every fixed $B \in \mathscr{F}_2$.* [4]

- *μ is uniformly σ-finite; that is, Ω_2 can be written as $\bigcup_n B_n$, where $B_n \in \mathscr{F}_2$ and there exist some constants $k_n \in \mathbb{R}^+$ such that $\mu(\omega_1, B_n) \le k_n$ for all $\omega_1 \in \Omega_1$.*

Then, there is a unique measure μ on $\mathscr{F} = \mathscr{F}_1 \times \mathscr{F}_2$ such that

$$\mu(A \times B) = \int_A \mu^{\omega_1}(B)\, d\mu_1 \text{ for every measurable rectangle } A \times B.$$

Namely,

$$\mu(C) = \int_{\Omega_1} \mu^{\omega_1}(C^{\omega_1})\, d\mu_1.$$

Furthermore, μ is σ-finite on $\mathscr{F}$; and if μ_1 and all the μ^{ω_1} are probability measures, so is μ.

Proof. Define the function $s(\omega_1, F) = \mu(\omega_1, F^{\omega_1})$.

1. First assume that the μ^{ω_1} are finite. We will start by showing that the function $s(\cdot, F)$ is measurable for every $C \in \mathscr{F}$.

   Let $\mathscr{C}$ be the class of sets of $\mathscr{F}$ for which the function is measurable. Note that $\mathscr{C}$ contains all measurable rectangles: take $C = A \times B$. Then,

   $$\mu^{\omega_1}(F^{\omega_1}) = \mu(\omega_1, F^{\omega_1}) = \begin{cases} \mu(\omega_1, A), & \text{if } \omega_1 \in A, \\ 0, & \text{otherwise.} \end{cases}$$

   Thus, $\mu^{\omega_1}(F^{\omega_1}) = \mu^{\omega_1}(A) I_A(\omega_1)$, which is clearly measurable.

   We will now use the Monotone Class Theorem to show that $s(\cdot, F)$ is measurable for all $F \in \mathscr{F}$. Let $\mathscr{F}_0$ be the field of disjoint unions of measurable rectangles.

   It is clear that $\mathscr{F}_0 \subseteq \mathscr{C}$, since, if $C_1, \ldots, C_n$ are disjoint measurable rectangles in $\mathscr{F}$, then for any given $\omega_1 \in \Omega_1$, from Proposition 4.1.2.(i) and Proposition 4.1.2.(ii) it is deduced that $\mu^{\omega_1}((\bigcup_k C_k)^{\omega_1}) = \sum_k \mu^{\omega_1}\left(C_k^{\omega_1}\right)$, which is a finite sum of Borel measurable functions; thus measurable.

   If we now show that $\mathscr{C}$ is a monotone class, it will follow that $\mathscr{F} = \sigma(\mathscr{F}_0) = \mathscr{C}$. If $A_n \uparrow A$, then $A_n^{\omega_1} \uparrow A^{\omega_1}$. By Proposition 2.2.11.(i), we have $\mu^{\omega_1}\left(A_n^{\omega_1}\right) \to \mu^{\omega_1}(A^{\omega_1})$; that is, $s(\omega_1, A_n^{\omega_1}) \to s(\omega_1, A^{\omega_1})$. Since a pointwise limit of measurable functions is measurable, we have $A^{\omega_1} \in \mathscr{C}$. Similarly, from the fact that all the measures are finite and 2.2.11.(ii), it follows that $\mathscr{C}$ also contains all descending limits of sets in it. Thus, $\mathscr{C} = \mathscr{F}$. This guarantees that μ is well-defined.

[4] Here it is convenient to remind the reader that, unless explicitly stated otherwise, the reals and extended reals are always equipped with the Borel σ-field.

Now it is easy to see that μ is a measure. Nonnegativity is immediate. Following Corollary 2.4.21.(i), and Proposition 4.1.2.(ii),

$$\begin{aligned}\mu\left(\bigcup_n A_n\right) &= \int_{\Omega_1} \mu^{\omega_1}\left(\left(\bigcup_n A_n\right)^{\omega_1}\right) d\mu_1\\ &= \int_{\Omega_1} \sum_n \mu^{\omega_1}\left(A_n^{\omega_1}\right) d\mu_1\\ &= \sum_n \int_{\Omega_1} \mu^{\omega_1}(A_n^{\omega_1})\, d\mu_1\\ &= \sum_n \mu(A_n).\end{aligned}$$

Finally, note that

$$\mu(A\times B) = \int_{\Omega_1} \mu^{\omega_1}((A\times B)^{\omega_1})\, d\mu_1 = \int_{\Omega_1} I_A(\omega_1)\mu^{\omega_1}(B)\, d\mu_1 = \int_A \mu^{\omega_1}(B)\, d\mu_1.$$

2. Now assume the μ^{ω_1} to be uniformly σ-finite. Note that the sets B_n may be assumed to be disjoint without loss of generality: simply take $B'_n = B_n \setminus \bigcup_{k=1}^{n-1} B_k$. Then, $\mu(\omega_1, B'_n) \le \mu(\omega_1, B_n) \le k_n$ for every ω_1.

   Define $\mu_n(\omega_1, B) = \mu(\omega_1, B\cap B_n)$ for each n. Note that $\mu^{\omega_1} = \sum_n \mu_n^{\omega_1}$ for every ω_1 (where $\mu_n^{\omega_1} = \mu_n(\omega_1,\cdot)$. Apply the construction above to each μ_n. Then, define a measure on $\mathscr{F}$ by
   $$\mu(C) = \sum_n \mu_n(C).$$

   This measure will satisfy

   $$\mu(C) = \sum_n \int_{\Omega_1} \mu_n^{\omega_1}(C^{\omega_1})\, d\mu_1 = \int_{\Omega_1} \sum_n \mu_n^{\omega_1}(C^{\omega_1})\, d\mu_1 = \int_{\Omega_1} \mu^{\omega_1}(C^{\omega_1})\, d\mu_1,$$

   which implies, again, that for measurable rectangles $\mu(A\times B) = \int_A \mu^{\omega_1}(B)\, d\mu_1$.

For the uniqueness part of the result, we will employ the Carathéodory Extension Theorem, by showing that μ is σ-finite on $\mathscr{F}_0$, the field of disjoint unions of measurable rectangles.

Since μ_1 is σ-finite, we can find a sequence of sets $A_1, A_2, \ldots$ such that $\mu_1(A_m) < \infty$ and $\Omega_1 = \bigcup_m A_m$. Define $C_{nm} = A_m \times B_n$, so that $\Omega = \bigcup_{nm} C_{nm}$. Additionally,

$$\mu(C_{nm}) = \int_{A_m} \mu^{\omega_1}(B_n)\, d\mu_1 \le k_n \mu_1(A_m) < \infty.$$

We have just seen that μ is σ-finite on $\mathscr{F}_0 \subseteq \mathscr{F}$, hence on $\mathscr{F}$. If μ_1 and all functions μ^{ω_1} are probability measures, it is immediate that μ is so too. □

A very natural question to ask now is how does integration with respect to this new measure relates with integration with respect to the "original" measures μ_1 and μ^{ω_1}. The following theorem answers with detail this question.

Theorem 4.1.5 (Two-Dimensional Fubini's Theorem). *Under the hypotheses and notation of Theorem 4.1.4, let* $\Omega = \Omega_1 \times \Omega_2$*, and* $f : \Omega \to \overline{\mathbb{R}}$ *be a* $\mathscr{F}$*-measurable function. For every fixed* $\omega_1 \in \Omega_1$*, consider the section* $f^{\omega_1} : \Omega_2 \to \overline{\mathbb{R}}$*. Then,*

(i) If f *is nonnegative, then the function* $F(\omega_1) = \int_{\Omega_2} f^{\omega_1} d\mu^{\omega_1}$ *exists and is* $\mathscr{F}_1$*-measurable. Also,*

$$\int_{\Omega} f\, d\mu = \int_{\Omega_1} F\, d\mu_1 = \int_{\Omega_1} \left(\int_{\Omega_2} f^{\omega_1}\, d\mu^{\omega_1} \right) d\mu_1 \tag{4.1}$$

(ii) If $\int_{\Omega} f\, d\mu$ *exists (respectively, is finite), then Equation* (4.1) *holds in the sense that the function* $F(\omega_1) = \int_{\Omega_2} f^{\omega_1}\, d\mu^{\omega_1}$ *exists (respectively, is finite) for* μ_1*-almost every* ω_1*, and defines a Borel measurable function of* ω_1 *if it is taken as* 0 *on the exceptional set (or as any Borel measurable function).*

Proof. 1. F exists because f^{ω_1} is nonnegative for every ω_1. Start by suppossing that f is an indicator, I_C. Then, it is not hard to see that $f^{\omega_1} = I_{C^{\omega_1}}$. Thus,

$$F(\omega_1) = \int_{\Omega_2} I_{C^{\omega_1}}\, d\mu^{\omega_1} = \int_{C^{\omega_1}} d\mu^{\omega_1} = \mu^{\omega_1}(C^{\omega_1}),$$

which is $\mathscr{F}_1$-measurable, as we saw in the proof of Theorem 4.1.4. Additionally, $\int_{\Omega} f\, d\mu = \mu(C) = \int_{\Omega_1} \mu^{\omega_1}(C^{\omega_1})\, d\mu_1 = \int_{\Omega_1} F(\omega_1)\, d\mu_1$, by that same theorem.

Now suppose f is a simple, nonnegative function $\sum_i x_i I_{C_i}$. Then, by linearity, $F(\omega_1) = \sum_i x_i \mu^{\omega_1}(C_i^{\omega_1})$, which is a measurable function and satisfies that

$$\int_{\Omega} f\, d\mu = \sum_i x_i \int_{\Omega} I_{C_i}\, d\mu = \int_{\Omega_1} F\, d\mu_1.$$

For an arbitrary nonnegative, Borel measurable function f, consider a sequence of nonnegative, simple functions $s_n \uparrow f$. Define $F_n(\omega_1) = \int_{\Omega_2} s_n^{\omega_1}\, d\mu^{\omega_1}$. By what we have seen for simple functions, F_n is measurable. By the Monotone Convergence Theorem, $F_n \uparrow F$ pointwise, and thus F is measurable. Applying the Monotone Convergence Theorem again, this time twice, it follows that

$$\int_{\Omega_1} F\, d\mu_1 = \lim_n \int_{\Omega_1} F_n\, d\mu_1 = \lim_n \int_{\Omega} s_n\, d\mu = \int_{\Omega} f\, d\mu.$$

2. Split $f = f^+ - f^-$. Note that $(f^{\omega_1})^+ = \left(f^+\right)^{\omega_1}$, and $(f^{\omega_1})^- = \left(f^-\right)^{\omega_1}$. Apply the construction above to f^+, f^-, and obtain two functions F^+, F^-. Note that $F(\omega_1)$ exists if, and only if, at least one of $F^+(\omega_1), F^-(\omega_1)$ is finite, and in that case $F(\omega_1) = F^+(\omega_1) - F^-(\omega_1)$.

Since $\int_{\Omega} f\, d\mu$ exists, at least one of f^+, f^- is integrable. Suppose, without loss of generality, that $\int_{\Omega} f^+\, d\mu < \infty$. Then,

$$\int_{\Omega_1} F^+\, d\mu_1 = \int_{\Omega} f\, d\mu < \infty,$$

by the first section of this theorem. Hence, F^+ is finite μ_1-a.e., and $F = F^+ - F^-$ exists whenever this is the case, and is measurable if taken as 0 (or any measurable function) elsewhere. Finally,

$$\int_\Omega f\,d\mu = \int_\Omega f^+\,d\mu - \int_\Omega f^-\,d\mu = \int_{\Omega_1} F^+\,d\mu_1 - \int_{\Omega_2} F^-\,d\mu_1 = \int_{\Omega_1} F\,d\mu_1.$$

The third integral is simply a different notation for the second integral. □

Remark 4.1.6. *In the last theorem, if we wish to integrate f in some measurable subset C of Ω, simply consider $g = I_C f$, which is measurable. Then, $\int_\Omega g\,d\mu$ exists if $\int_\Omega f\,d\mu$ does and $g^{\omega_1} = I_{C^{\omega_1}} f^{\omega_1}$. Thus,*

$$\int_C f\,d\mu = \int_\Omega g\,d\mu_1 = \int_{\Omega_2}\int_{C^{\omega_1}} f^{\omega_1}\,d\mu^{\omega_1}\,d\mu_1.$$

Having covered the $n = 2$ case, we can now inductively extend the last two theorems to products of n measurable spaces.

Theorem 4.1.7. *Let $(\Omega_j, \mathscr{F}_j)$ be measurable spaces, $j = 1, \ldots, n$, and μ_1 a σ-finite measure on $\mathscr{F}_1$. Assume that, for each $j = 1, \ldots, n-1$, we are given a map $\mu_{j+1} \colon \Omega_1 \times \cdots \times \Omega_j \times \mathscr{F}_{j+1} \to \overline{\mathbb{R}}$ that satisfies*

- *The mapping $\mu^{\omega_1,\ldots,\omega_j} = \mu_{j+1}(\omega_1, \ldots, \omega_j, \cdot)$ is a measure over $\mathscr{F}_{j+1}$ for every fixed $\omega_1 \in \Omega_1, \ldots, \omega_j \in \Omega_j$.*
- *$\mu_{j+1}(\cdot, \ldots, \cdot, B)$ is $(\mathscr{F}_1 \times \cdots \times \mathscr{F}_j)$-measurable for every fixed $B \in \mathscr{F}_{j+1}$.*
- *μ_{j+1} is uniformly σ-finite, that is, Ω_{j+1} can be written as $\bigcup_r B_{(j+1)r}$, where $B_{(j+1)r} \in \mathscr{F}_{j+1}$ and there exist some constants $k_n \in \mathbb{R}^+$ such that $\mu_{j+1}(\omega_1, \ldots, \omega_j, B_n) \le k_n$ for all $\omega_1 \in \Omega_1, \ldots, \omega_j \in \Omega_j$.*

Then,

(i) **Product Measure Theorem.** *There exists a unique measure μ on $\mathscr{F} = \mathscr{F}_1 \times \cdots \times \mathscr{F}_n$ such that*

$$\mu(A_1 \times \cdots \times A_n) = \int_{A_1}\left(\int_{A_2}\left(\cdots\left(\int_{A_n} d\mu^{\omega_1,\ldots,\omega_{n-1}}\right)\cdots\right)d\mu^{\omega_1}\right)d\mu_1$$

for every measurable rectangle $A_1 \times \cdots \times A_n$.

(ii) **Fubini's Theorem.** *Let $\Omega = \Omega_2 \times \cdots \times \Omega_n$ and $f \colon \Omega \to \overline{\mathbb{R}}$ be a $\mathscr{F}$-measurable function. For every fixed $\omega_1 \in \Omega_1, \ldots, \omega_{n-1} \in \Omega_{n-1}$, consider the section $f^{\omega_1,\ldots,\omega_{n-1}}(\omega_n) = f(\omega_1, \ldots, \omega_n)$. Then,*

a) If f is nonnegative,

$$\int_\Omega f\,d\mu = \int_{A_1}\left(\int_{A_2}\left(\cdots\left(\int_{A_n} f^{\omega_1,\ldots,\omega_{n-1}}\,d\mu^{\omega_1,\ldots,\omega_{n-1}}\right)\cdots\right)d\mu^{\omega_1}\right)d\mu_1, \tag{4.2}$$

where, after the integration with respect to each $\mu^{\omega_1,\ldots,\omega_j}$, the result is a $(\mathscr{F}_1 \times \cdots \times \mathscr{F}_j)$-measurable function of $\omega_1, \ldots, \omega_j$.

b) *If $\int_\Omega f\ d\mu$ exists (respectively, is finite), Equation* (4.2) *holds in the sense that, for each $j = n-1,\dots,1$, the integral with respect to $\mu^{\omega_1,\dots,\omega_j}$ exists (is finite) λ_j-almost everywhere, where $\lambda_1 = \mu_1$ and, for $j \geq 2$, λ_j is the measure on $\mathscr{F}_1 \times \cdots \times \mathscr{F}_j$ determined by the measures $\mu_1, \mu^{\omega_1}, \dots, \mu^{\omega_1,\dots,\omega^{j-1}}$ as in* (a)*; additionally, if the integral is defined as* 0 *(or as any other jointly measurable function) on the exceptional set, then it is a $\mathscr{F}_1 \times \cdots \times \mathscr{F}_j$-measurable function.*

Proof. We will prove both sections of the theorem together, by induction on n.
For the base case, $n = 2$, 4.1.7.(i) is simply Theorem 4.1.4, where we simply rewrote $\mu^{\omega_1}(A_2) = \int_{A_2} d\mu^{\omega_1}$. Note that 4.1.7.(ii) for $n = 2$ is exactly Theorem 4.1.5.
For the inductive step, suppose the result holds for the first $n-1$ factors, and obtain a σ-finite measure λ_{n-1} on $\mathscr{F}_1 \times \cdots \times \mathscr{F}_{n-1}$ using 4.1.7.(i) and the measures $\mu^{\omega_{i_1},\dots,\omega_{i_{n-1}}}$.
For 4.1.7.(i), we need to construct a measure on $\mathscr{F}_1 \times \cdots \times \mathscr{F}_n$. It is not difficult to prove that, if we identify every cartesian product $(\Omega_1 \times \Omega_2) \times \Omega_3$ with its *associative conjugate* $\Omega_1 \times (\Omega_2 \times \Omega_3)$, then product measure spaces are also associative, in the sense that $\mathscr{F}_1 \times \cdots \times \mathscr{F}_n = (\mathscr{F}_1 \times \cdots \times \mathscr{F}_{n-1}) \times \mathscr{F}_n$. Thus, applying the base case $n = 2$ to $\Omega_1 \times \cdots \times \Omega_{n-1}$ and Ω_n, there exists a unique measure μ on $\mathscr{F}_1 \times \cdots \times \mathscr{F}_n$ such that

$$\mu(P \times A_n) = \int_P \left(\int_{A_n} d\mu^{\omega_1,\dots,\omega_{n-1}} \right) d\lambda_{n-1}$$

for every measurable rectangle $P \times A_n \in (\mathscr{F}_1 \times \cdots \times \mathscr{F}_{n-1}) \times \mathscr{F}_n$. In the case that P itself is a measurable rectangle $A_1 \times \cdots \times A_{n-1}$, we have $I_P(\omega_1,\dots,\omega_{n-1}) = I_{A_1}(\omega_1)\dots I_{A_{n-1}}(\omega_{n-1})$, and the expression becomes $\mu(A_1 \times \cdots \times A_n) = \int_{\Omega_1 \times \cdots \times \Omega_{n-1}} g\ d\lambda_{n-1}$, where $g(\omega_1,\dots,\omega_{n-1}) = I_{A_1}(\omega_1)\dots I_{A_{n-1}}(\omega_{n-1}) \int_{A_n} d\mu^{\omega_1,\dots,\omega_{n-1}}$. Applying the induction hypothesis in 4.1.7.(ii) to g, and factoring out the indicator functions I_{A_j} to the left-most possible, we have

$$\mu(A_1 \times \cdots \times A_n) = \int_{\Omega_1} I_{A_1} \left(\int_{\Omega_2} I_{A_2} \left(\cdots \left(\int_{A_n} d\mu^{\omega_1,\dots,\omega_{n-1}} \right) \cdots \right) d\mu^{\omega_1} \right) d\mu_1.$$

This is the desired existence result in 4.1.7.(i). To show uniqueness, it suffices to show that the condition imposed to μ implies that it is σ-finite on the field of disjoint unions of measurable rectangles. For every $j = 1,\dots,n$, write $\Omega_j = \bigcup_r B_{jr}$ so that μ_j is uniformly σ-finite on the sets B_{jr} with bounds k_{jr}. Then, write $\Omega = \bigcup_{j,r} B_{jr}$, so that

$$\mu(B_{jr}) \leq k_{1r}\dots k_{nr} < \infty.$$

This finishes the proof of 4.1.7.(i). To prove 4.1.7.(ii), simply note that the measure μ obtained in 4.1.7.(i) was obtained using the $n = 2$ case with λ_{n-1} and the measures $\mu^{\omega_{i_1},\dots,\omega_{i_{n-1}}}$. Thus, by the two-dimensional Fubini's Theorem,

$$\int_\Omega f\ d\mu = \int_{\Omega_1 \times \cdots \times \Omega_{n-1}} \left(\int_{\Omega_n} f^{\omega_1,\dots,\omega_{n-1}}\ d\mu^{\omega_1,\dots,\omega_{n-1}} \right) d\lambda_{n-1},$$

where the inner integral is Borel measurable in $\mathscr{F}_1 \times \cdots \times \mathscr{F}_{n-1}$, or becomes so after redefining it on a λ_{n-1}-null set. The rest of the result is a straightforward consequence of the induction hypothesis on 4.1.7.(ii). □

We finish the section with two special cases of the theorem proved above, which the reader versed in analysis might be slightly more familiar with.

Corollary 4.1.8 (Classical Product Measure Theorem). *Let $\left(\Omega_k, \mathscr{F}_k, \mu_k\right)$, $k = 1, \dots, n$, be σ-finite measure spaces. Then, there exists a unique measure μ on $\mathscr{F} = \mathscr{F}_1 \times \cdots \times \mathscr{F}_n$, called the **product measure** of $\mu_1, \dots, \mu_n$, such that*

$$\mu(A_1 \times \cdots \times A_n) = \mu_1(A_1) \dots \mu_n(A_n)$$

on every measurable rectangle $A_1 \times \cdots \times A_n$.

Proof. Take $\mu_{j+1}(\omega_1, \dots, \omega_j, B) = \mu_{j+1}(B)$ for every $j = 1, \dots, n-1$ in Theorem 4.1.7.(i). Then, all hypotheses are satisfied, and every section $\mu^{\omega_1, \dots, \omega_j} = \mu_{j+1}$, hence there exists a unique measure μ on $\mathscr{F} = \mathscr{F}_1 \times \cdots \times \mathscr{F}_n$ such that

$$\mu(A_1 \times \cdots \times A_n) = \int_{A_1} \left(\int_{A_2} \left(\cdots \left(\int_{A_n} d\mu_n \right) \cdots \right) d\mu_2 \right) d\mu_1 = \mu_1(A_1) \dots \mu_n(A_n)$$

for every measurable rectangle $A_1 \times \cdots \times A_n$. □

Finally, a simple consequence of the construction realised so far is a very useful theorem in multivariate calculus.

Corollary 4.1.9 (Classical Fubini's Theorem). *Following the notation and hypotheses of Corollary 4.1.8, write $\Omega = \Omega_1 \times \cdots \times \Omega_n$. If $f \colon \Omega \to \overline{\mathbb{R}}$ is a $\mathscr{F}$-measurable function such that $f \geq 0$ or $\int_\Omega f \, d\mu$ exists, then*

$$\int_\Omega f \, d\mu = \int_{\Omega_1} \int_{\Omega_2} \cdots \int_{\Omega_n} f \, d\mu_n \dots d\mu_2 d\mu_1. \tag{4.3}$$

By symmetry, the integration can be performed in any order. Additionally, if the iterated integral of $|f|$ in any specific order is finite, then so is $\int_\Omega |f| \, d\mu$; hence, $\int_\Omega f \, d\mu$ exists and Equation (4.3) *holds; in particular, the integration can be performed in any order too.*

4.2 Infinite product of probabilities

It is possible to extend the construction made for finite product spaces to countable products without running into too much trouble, except that it becomes convenient to impose that the measures are finite. In this text, however, we will skip straight to arbitrary products. The interested reader can see Section 2.7 of [6].

For arbitrary infinite products, the measures are required to be finite too, and we additionally require some topological properties on the base sets which boil down to being able to take profit of sequential compactness.

We begin the section introducing some notation:

Definition 4.2.1. *Let $\{\Omega_t\}_{t \in T}$ be an arbitrarily indexed family of nonempty sets. Let $\prod_t \Omega_t$ be the **product** of the sets Ω_t, that is, the set of all functions $\omega \colon T \to \bigcup_t \Omega_t$ such that $\omega(t) \in \Omega_t$ for all t*[5]*. We will write $\Omega = \prod_t \Omega_t$. Just to fix notation, suppose that the*

[5] The axiom of choice is equivalent to the statement that $\prod_t \Omega_t$ is nonempty for every family of nonempty sets $\{\Omega_t\}_{t \in T}$.

index set T is totally ordered[6]*, and consider a finite subset $v \subseteq T$. Write $v = \{t_1, \ldots, t_n\}$, with $t_1 < \ldots < t_n$. The finite product $\prod_{i=1}^{n} \Omega_{t_i}$ is denoted by Ω_v. Similarly, if $\omega \in \Omega$, the notation ω_v will be used for the tuple $(\omega(t_1), \ldots, \omega(t_n)) \in \Omega_v$. For every $B \subseteq \Omega_v$, we define the* ***cylinder with base B*** *as*

$$B_v = \{\omega \in \Omega | \omega_v \in B\}.$$

Remark 4.2.2. *Cylinders have interesting algebraic properties regarding set operations, in the sense that, for any $B \subseteq \Omega_v$ and any family of sets $B_i \subseteq \Omega_v$, $i \in I$ (I is some arbitrary index set), we have*

$$(B_v)^c = \left(B^c\right)_v, \quad \bigcup_{i \in I} (B_i)_v = \left(\bigcup_{i \in I} B_i\right)_v \quad \text{and} \quad \bigcap_{i \in I} (B_i)_v = \left(\bigcap_{i \in I} B_i\right)_v$$

Definition 4.2.3. *Let $v = \{t_1, \ldots, t_n\}$, $t_1 < \ldots < t_n$ be a finite subset of T and $u = \left\{t_{i_1} \ldots t_{i_k}\right\}$, with $t_{i_1} < \ldots < t_{i_k}$, a nonempty subset of v. If we consider some $y = \left(y_{t_1}, \ldots, y_{t_n}\right) \in \Omega_v$, then the symbol y_u will be used to denote the k-tuple $(y_{t_{i_1}}, \ldots, y_{t_{i_k}})$. This notation is consistent with the one introduced for Ω in the sense that, for any $\omega \in \Omega$, one has $(\omega_v)_u = \omega_u$. Also, if $B \subseteq \Omega_u$, we define the* ***retraction***[7] *of B in Ω_v as*

$$B^v = \{\omega \in \Omega_v | \omega_u \in B\}.$$

Remark 4.2.4. *Just as cylinders, retractions have interesting algebraic properties regarding set operations: if $B, B_i \subseteq \Omega_u$, $i \in I$ for some arbitrary index set I, then*

$$\left(B^v\right)^c = \left(B^c\right)^v, \quad \bigcup_{i \in I} (B_i)^v = \left(\bigcup_{i \in I} B_i\right)^v \quad \text{and} \quad \bigcap_{i \in I} (B_i)^v = \left(\bigcap_{i \in I} B_i\right)^v$$

Now that we have established a notation, the idea is to construct a σ-field on the product space Ω so that it is comfortable to take the leap from "finite" subsets $B \subseteq \Omega_v$ to the infinite product space Ω. Cylinders provide a good way to do this, so we will choose our σ-field so that it is easy to work with them.

Definition 4.2.5. *Continuing with the notation introduced in Definition 4.2.1, now suppose that every Ω_t is equipped with a σ-field $\mathscr{F}_t$. The symbol $\mathscr{F}_v$ is used to denote the (finite) product σ-field $\mathscr{F}_{t_1} \times \cdots \times \mathscr{F}_{t_n}$. We say that a cylinder B_v is a* ***measurable cylinder*** *whenever $B \in \mathscr{F}_v$, and a* ***measurable rectangle*** *whenever $B = B_1 \times \cdots \times B_n$, with $B_i \in \mathscr{F}_{t_i}$.*

Any (measurable) cylinder can be regarded as having a higher-dimensional base, in the sense that, if w is another finite subset of T such that $v \subseteq w$, then $B_v = (B^w)_w$. It is then not hard to see that measurable cylinders form a field (we can always regard a finite number of cylinders as having "the same dimension"). The σ-field generated by the field of measurable cylinders is called the ***product*** *of the σ-fields $\mathscr{F}_t$, and denoted by $\prod_t \mathscr{F}_t$. We will also write $\mathscr{F} = \prod_t \mathscr{F}_t$ if no confusion arises.*

[6] If T is a subset of $\mathbb{R}$, which will be the most common case, this is immediate. As a curiosity, another equivalence of the Axiom of Choice is that every set can be (somehow) well-ordered. We will not be using the ordering of T for anything other than notation, so this condition is nonrestrictive.

[7] Although a similar notation is used for convenience, retractions and sections are **not** the same concept when the product is finite.

Remark 4.2.6. *Remark 4.2.4 allows us to conclude that retractions are always measurable in their respective σ-fields: define $\mathcal{M} = \{B \in \mathscr{F}_u \mid B^v \in \mathscr{F}_v\}$. It is then clear that $\mathcal{M}$ is a σ-field over Ω_u that contains all measurable rectangles in $\mathscr{F}_u$, hence $\mathcal{M} = \mathscr{F}_u$.*

Having constructed the "framework" σ-field $\mathscr{F} = \prod_{t\in T} \mathscr{F}_t$ and developed the necessary language, we now wish to construct measures on it. Because we are working with infinite products, the construction will run into trouble for non-finite measures, hence we will restrict ourselves to working with probability measures. Suppose that we already have a probability measure P on $\mathscr{F}$. Note that P induces new probability measures in all of the projection spaces: consider some finite subset $v \subseteq T$. We may then define a probability measure P_v on $\mathscr{F}_v$, called the **projection** of P to $\mathscr{F}_v$, as

$$P_v(B) = P(B_v).$$

This is a probability measure because P is and Remark 4.2.2.

This new family of probability measures P_v, where v ranges over the finite subsets of T, satisfies a property called **consistency**: if v is a finite subset of T, and u is a nonempty subset of v, then the process described in Definition 4.2.5 of "regarding cylinders as having a higher dimension" leaves their measures unchanged, in the sense that for any $B \in \mathscr{F}_u$,

$$P_u(B) = P(B_u) = P\left((B^v)_v\right) = P_v(B^v).$$

We are able to construct measures somewhat easily in finite product spaces, so a very interesting question to be made now is if we are able to "reconstruct" the probability measure P just from the projections on "finite subspaces" $\mathscr{F}_v$. The Kolmogorov Extension Theorem (also called sometimes the Kolmogorov Consistency Theorem) gives a positive answer to this question, provided certain topological assumptions are made over the base measurable spaces $(\Omega_t, \mathscr{F}_t)$. Before proceeding with the final theorem of this work, we need a preliminary discussion regarding the topological properties of the base spaces.

Suppose that we have a finite family of metric spaces, $\Omega_1, \ldots, \Omega_n$, where each Ω_k is equipped with a metric d_k. Write $v = \{1, \ldots, n\}$, and denote by Ω_v the (topological) product space $\Omega_1 \times \cdots \times \Omega_n$. It is now an elementary result in topology that Ω_v is metrizable when equipped with the euclidean distance

$$d(x, y) = \left(\sum_{k=1}^{n} d_k(x_k, y_k)^2 \right)^{\frac{1}{2}},$$

where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$. Moreover, convergence in Ω_v is characterised by convergence on each "component" space Ω_k, in the sense that a sequence $\{x^m\}_{m\in\mathbb{Z}^+}$ is Cauchy in Ω_v if, and only if, each of the component sequences $\{x_k^m\}_{m\in\mathbb{Z}^+}$ is Cauchy in its respective metric space Ω_k, and $x^m \to x \in \Omega_v$ if, and only if, $x_k^m \to x_k \in \Omega_k$ for every $k = 1, \ldots, n$. Taking this into account, the following result is immediate:

Lemma 4.2.7. *Let $n \in \mathbb{Z}^+$, and suppose that for every $k = 1, \ldots, n$ we are given a metric space (Ω_k, d_k). Write $v = \{1, \ldots, n\}$, and consider the product space $\Omega_v = \Omega_1 \times \cdots \times \Omega_n$. Then, Ω_v is a complete metric space if, and only if, each of the spaces Ω_k is complete.*

Moreover, if u is a nonempty subset of v, write $u = \{i_1, \ldots, i_l\}$, with $i_1 < \ldots < i_l$. Now define $\Omega_u = \Omega_{i_1} \times \cdots \times \Omega_{i_l}$, and for any given $x = (x_1, \ldots, x_n) \in \Omega_v$, define $x_u = (x_{i_1}, \ldots, x_{i_l})$. Then, if $\{x^m\}_{m \in \mathbb{Z}^+}$ is a sequence in Ω_v and $x \in \Omega_v$, we have that

$$x^m \to x \implies \left(x^m\right)_u \to x_u.$$

Finally, one last result is required:

Lemma 4.2.8. *Continuing with the notation of Lemma 4.2.7, now suppose that each Ω_k is separable and equipped with the σ-field $\mathscr{F}_k = \mathscr{B}(\Omega_k)$. Then, Ω_v is separable and the product σ-field induced coincides with the Borel sets of the product topological space: namely, $\mathscr{F}_1 \times \cdots \times \mathscr{F}_n = \mathscr{B}(\Omega_v)$.*[8]

Proof. By Lemma 3.2.12, for each k there exists a countable basis $\left\{B_{km} \middle| m \in \mathbb{Z}^+\right\}$ for Ω_k. It is then clear that $\left\{B_{1i_1} \times \cdots \times B_{ni_n} \middle| m_1, \ldots, m_n \in \mathbb{Z}^+\right\}$ forms a countable basis for Ω_v, hence it is separable.
Call $\mathscr{F}_v = \mathscr{F}_1 \times \cdots \times \mathscr{F}_n$. Then, each of the sets $B_{1n_1} \times \cdots \times B_{kn_k}$ is a measurable rectangle, hence it is in $\mathscr{F}_v$. Since every open set is a countable union of sets of this form, it follows that every open set is in $\mathscr{F}_v$, thus $\mathscr{B}(\Omega_v) \subseteq \mathscr{F}_v$.
For the reciprocal inclusion, fix some $k = 1, \ldots, n$. Define

$$\mathscr{M}_k = \left\{B \in \mathscr{F}_k \middle| \{\omega \in \Omega_v : \omega_k \in B\} \in \mathscr{B}(\Omega_v)\right\}$$

Now, if B is open, $\{\omega \in \Omega_v : \omega_k \in B\} = \Omega_1 \times \cdots \times \Omega_{k-1} \times B \times \Omega_{k+1} \times \cdots \times \Omega_n$ is open, hence in $\mathscr{B}(\Omega_v)$. It follows that $\mathscr{M}_k$ is a σ-field containing all open sets in Ω_k, and therefore $\mathscr{M}_k = \mathscr{F}_k$. Now consider a measurable rectangle $B_1 \times \cdots \times B_n$. From this,

$$B_1 \times \cdots \times B_n = \bigcap_{k=1}^{n} \{\omega \in \Omega_v : \omega_k \in B_k\} \in \mathscr{B}(\Omega_v),$$

and therefore $\mathscr{F}_v \subseteq \mathscr{B}(\Omega_v)$. □

We are finally in position to enunciate and prove the main result.

Theorem 4.2.9 (Kolmogorov Extension Theorem). *Let T be an arbitrary index set, and for every $t \in T$, let Ω_t be a complete, separable metric space. Let $\mathscr{F}_t$ be the Borel sets of Ω_t. Assume that for each finite, nonempty subset $v \subseteq T$ we are given a probability measure P_v on $\mathscr{F}_v$. Assume that this family of measures is consistent, that is, for each nonempty subset $u \subseteq v$ and $B \in \mathscr{F}_u$, we have*

$$P_u(B) = P_v\left(B^v\right).$$

Then, there exists a unique probability measure P on $\mathscr{F} = \prod_t \mathscr{F}_t$ such that

$$P_v(A) = P(A_v)$$

for all v and every $A \in \mathscr{F}_v$.

[8] Both this result and the previous one are original and they deviate slightly from the contents of [6]. There, a similar result is stated for countable products; however, no countable products are then used in the proof of the Kolmogorov Extension Theorem (only finitely many at a time). Also, no indication is given as to how the finite projection spaces relate to the countable product.

Proof. Define the desired probability measure on measurable cylinders by $P(C) = P_v(B)$, where v is some finite, nonempty subset of T and B is some set in $\mathscr{F}_v$ such that $C = B_v$. The definition of a measurable cylinder guarantees that there exists at least one such set B, and the consistency property guarantees that this does not depend explicitly on B.

Let $\mathscr{F}_0$ be the field of measurable cylinders. We know that $\mathscr{F} = \sigma(\mathscr{F}_0)$, so we will use the Carathéodory Extension Theorem to extend P to $\mathscr{F}$. To see that P is additive on $\mathscr{F}_0$, consider disjoint measurable cylinders $C_1, \dots, C_n$. Write each $C_k = (B_k)_{v_k}$, where v_k are nonempty subsets $v_1, \dots, v_n \subseteq T$ and with $B_k \in \mathscr{F}_{v_k}$. Let $v = \bigcup_k v_k$, so that $C_k = ((B_k)^v)_v$. Then,

$$P\left(\bigcup_k C_k\right) = P\left(\left(\bigcup_k B_k^v\right)_v\right) = P_v\left(\bigcup_k B_k^v\right) = \sum_k P_v\left(B_k^v\right) = \sum_k P(C_k).$$

To see that P is σ-additive on $\mathscr{F}_0$, the idea is to show that it is continuous from above at $\emptyset$ and use Proposition 2.2.12.(ii). Then, let A_k be a sequence of sets in $\mathscr{F}_0$ decreasing to $\emptyset$. Since P is additive, it follows that $P(A_k)$ is a decreasing sequence. Suppose that it does not approach 0, that is, there exists some $\varepsilon > 0$ such that $P(A_k) \geq \varepsilon$ for all k. Write $A_k = (D_k)_{v_k}$, where $D_k \in \mathscr{F}_{v_k}$. By regarding the bases as having a higher dimension, we may assume, without loss of generality, that the sets v_k increase with k.

By Lemmas 4.2.7 and 4.2.8, each Ω_{v_k} is a separable, complete metric space and $\mathscr{F}_{v_k} = \mathscr{B}\left(\Omega_{v_k}\right)$. Therefore, by Theorem 3.2.11, there exists some compact set $C_k \subseteq D_k \subseteq \Omega_{v_k}$ such that $P_{v_k}(D_k \setminus C_k) < \varepsilon 2^{-k-1}$. Define $A'_k = (C_k)_{v_k} \in \mathscr{F}$, so that $A'_k \subseteq A_k$ and $P\left(A_k \setminus A'_k\right) < \varepsilon 2^{-k-1}$. In this way, we can approximate the given cylinders by cylinders with compact bases.

Now define $F_k = \bigcap_{i=1}^k A'_i \subseteq \bigcap_{i=1}^k A_i = A_k$. Then,

$$P(A_k \setminus F_k) = P\left(\bigcup_{i=1}^k (A_k \setminus A'_i)\right) \leq \sum_{i=1}^k P\left(A_k \setminus A'_i\right) \leq \sum_{i=1}^k P(A_i \setminus A'_i) < \sum_{i=1}^k \varepsilon 2^{-i-1} < \frac{\varepsilon}{2}.$$

Since $P(A_k \setminus F_k) = P(A_k) - P(F_k) < \frac{\varepsilon}{2}$, we have that $P(F_k) > P(A_k) - \frac{\varepsilon}{2} \geq \frac{\varepsilon}{2} > 0$. In particular, F_k is nonempty.

Now pick some $x^k \in F_k$ for each k. Consider the sequence $x^1_{v_1}, x^2_{v_1}, x^3_{v_1}, \dots$ Since the $x^n_{v_1}$ belong to C_1, a compact subset of a metric space (thus, sequentially compact), there exists a subsequence $x^{r_{1n}}_{v_1}$ approaching some $x_{v_1} \in C_1$ when $n \to +\infty$. Now consider the sequence $x^{r_{11}}_{v_2}, x^{r_{12}}_{v_2}, x^{r_{13}}_{v_2}, \dots$ Since $x^k_{v_2} \in C_2$ for every $k \geq 2$, it follows that, eventually, the sequence considered is in C_2. Again, there exists a subsequence, $x^{r_{2n}}_{v_2}$ approaching some $x_{v_2} \in C_2$. Note that, for every n, $\left(x^{r_{2n}}_{v_2}\right)_{v_1} = x^{r_{2n}}_{v_1}$. As $n \to +\infty$, the left side approaches $\left(x_{v_2}\right)_{v_1}$ (Lemma 4.2.7 may be used to see this), and since r_{2n} is a subsequence of r_{1n}, the right side approaches x_{v_1}; hence $\left(x_{v_2}\right)_{v_1} = x_{v_1}$. Continue in this fashion and construct a sequence x_{v_i} in such a way that, at step i, we have

$$x^{r_{in}}_{v_i} \to x_{v_i} \in C_i \text{ as } n \to +\infty, \quad \text{and} \quad \left(x_{v_i}\right)_{v_j} = x_{v_j} \text{ for all } j < i.$$

Now choose some $\omega \in \Omega$ such that $\omega_{v_i} = x_{v_i}$ for all $i \in \mathbb{Z}^+$ (remember that the elements x_{v_i} are simply finite tuples; hence such a choice is possible because $\left(x_{v_i}\right)_{v_j} = x_{v_j}$ for $j < i$). Then $\omega_{v_i} \in C_i$ for each i. Therefore,

$$\omega \in \bigcap_i A'_i \subseteq \bigcap_i A_i = \emptyset,$$

a contradiction. It must be that $P(A_k) \to 0$. Therefore, P is σ-additive on $\mathscr{F}_0$. Since it is trivially σ-finite for being bounded by 1 (every P_v is a probability measure), it extends to a probability measure P on $\mathscr{F}$.

If Q is another such probability measure on $\mathscr{F}$, then it agrees with P in measurable cylinders:

$$Q(B_v) = P_v(B) = P(B_v),$$

hence $Q = P$ by the uniqueness part of the Carathéodory Extension Theorem. □

CHAPTER 5

AN APPLICATION TO REAL ANALYSIS AND PROBABILITY THEORY

Throughout this text, we have developed some powerful tools in Measure Theory. Some of them will allow us to gain some insight into Real Analysis, and others - specially those developed in the last chapter, regarding product spaces - will allow us to answer some questions in Probability Theory with respect to the existence of certain objects.

5.1 Some applications to Real Analysis

So far, we have developed all of Measure Theory *abstractly*, that is, without mentioning any concrete measures. In Analysis, however, we are mostly interested in *classical* notions, such as areas and volumes. The Riemann integral is, precisely, a formalisation and generalisation of the notion of area under a curve. Luckily, there is a concrete measure on $\mathbb{R}^n$ that achieves this: the **Lebesgue measure**.

We will begin by working on $\mathbb{R}$. Here, the most basic metric notion is that of the *longitude* of a segment. If we want to construct a measure that represents this and is also able to give values to wide enough class of sets, we could define it on $\mathscr{B}(\mathbb{R})$ and it should assign its length to each interval.

The approach followed here will allow us to construct a broad class of measures on $\mathscr{B}(\mathbb{R})$, which will be useful when working in Probability Theory. Mainly, we will study two concepts:

Definition 5.1.1. *A* ***Lebesgue-Stieltjes*** *measure on $\mathbb{R}$ is a measure μ over the usual σ-field $\mathscr{B}(\mathbb{R})$ such that $\mu(I) < \infty$ for every bounded interval I. A* ***distribution function*** *on $\mathbb{R}$ is a mapping $F\colon \mathbb{R} \to \mathbb{R}$ that is* ***increasing*** *($a \le b$ implies $F(a) \le F(b)$) and* ***right-continuous*** *(that is, $\lim_{x\to x_0^+} F(x) = F(x_0)$[1]).*

[1] Equivalently, $\lim_n F(x_n) = F(x_0)$ for every sequence $x_n \downarrow x_0$.

There exists a close relation between the concepts defined above. Namely, the formula $\mu((a,b]) = F(b) - F(a)$ yields a one-to-one correspondence between Lebesgue-Stieltjes measures and distribution functions, up to an additive constant.

If provided with the measure μ, checking that F is a distribution function is very simple using the properties of measures developed through Section 2.2.

Proposition 5.1.2. *Let μ be a Lebesgue-Stieltjes measure on $\mathbb{R}$. Define $F(0)$ as any real number. Then, the function $F\colon \mathbb{R} \to \mathbb{R}$ defined as*

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

is a distribution function.

Reciprocally, if provided with the distribution function F, constructing the associated Lebesgue-Stieltjes measure is slightly harder. A sketch of the proof is given below:

Theorem 5.1.3. *Let $F\colon \mathbb{R} \to \mathbb{R}$ be a distribution function. Then, there exists a unique measure on $\mathscr{B}(\mathbb{R})$ that satisfies the formula $\mu((a,b]) = F(b) - F(a)$, $a < b \in \mathbb{R}$. Moreover, this is a Lebesgue-Stieltjes measure.*

Proof. The idea is to find a suitable σ-field to use the Carathéodory Extension Theorem. In $\overline{\mathbb{R}}$, the class of disjount unions of right-semiclosed intervals, $\mathscr{F}_0(\overline{\mathbb{R}})$, forms a field. Also, since $\overline{\mathbb{R}}$ is compact (one way to see this is by noting that it is homeomorphic to $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ via $\arctan(x)$), it will be more convenient to work in $\overline{\mathbb{R}}$ than in $\mathbb{R}$. Extend F to $\overline{\mathbb{R}}$ as $F(-\infty) = \lim_{x\to-\infty}$and $F(+\infty) = \lim_{x\to+\infty} F(x)$ (both limits exist by monotonicity).

Now μ is defined on right-semiclosed intervals as $\mu((a.b]) = F(b) - F(a)$ (where $a \leq b \in \overline{\mathbb{R}}$) and $\mu([-\infty, a]) = F(a) - F(-\infty)$. Then μ is extended to $\mathscr{F}_0(\overline{\mathbb{R}})$ by additivity. It is now shown that μ is countably additive by using the compactness of $\overline{\mathbb{R}}$ and Proposition 2.2.12.(i) (this is Lemma 1.4.3 of [6]):

Let $A_1, A_2, \dots$ be a sequence of sets in $\mathscr{F}_0(\overline{\mathbb{R}})$ decreasing to $\emptyset$. Note that each set A_n is the union of a finite number of disjoint right-semiclosed intervals, and $\mu\left(a', b\right] \to \mu(a,b]$ when $a' \to a^+$ for each $a \leq b \in \overline{\mathbb{R}}$. Then, for every $\varepsilon > 0$, it is possible to find a set B_n such that $\overline{B}_n \subseteq A_n$ and $\mu\left(A_n \setminus \overline{B}_n\right) < \varepsilon 2^{-n}$. Note that

$$\bigcap_n \overline{B}_n \subseteq \bigcap_n A_n = \emptyset,$$

hence the sets $\overline{B}_n^c$ form an open covering of the compact $\overline{\mathbb{R}}$. Therefore, there exists a finite collection $\overline{B}_{n_1}, \dots, \overline{B}_{n_k}$ such that $\bigcap_{i=1}^k \overline{B}_{n_i} = \emptyset$. Take $n_0 = \max(n_1, \dots, n_k)$. Therefore, for each $n \geq n_0$, we have $A_n \subseteq A_{n_i}$ for every $i = 1, \dots, k$. Thus,

$$\mu(A_n) = \mu\left(A_n \setminus \left(\bigcup_{i=1}^k \overline{B}_{n_i}\right)\right) + \mu\left(\bigcap_{i=1}^k \overline{B}_{n_i}\right) \leq \mu\left(\bigcup_{i=1}^k \left(A_n \setminus \overline{B}_{n_i}\right)\right) + 0 \leq \mu\left(\bigcup_{i=1}^k \left(A_{n_i} \setminus \overline{B}_{n_i}\right)\right) < \varepsilon,$$

where in the last step we used that $\mu\left(\bigcup_n C_n\right) \leq \sum_n \mu(C_n)$ for every sequence of measurable sets C_n whose union is measurable too.

If μ were σ-finite on $\mathscr{F}_0(\overline{\mathbb{R}})$, we could extend it to a σ-field, but this need not be the case in general. This may be solved by restricting μ to $\mathscr{F}_0(\mathbb{R})$, the field of disjoint unions of right-semiclosed intervals (counting $(a,+\infty)$ as right-semiclosed): here, we may consider the sets $(-n,n]$, which cover $\mathbb{R}$ and on which μ is clearly finite.

We have defined μ on the desired field and checked the necessary hypotheses, and hence it extends to a unique measure on $\sigma(\mathscr{F}(\mathbb{R})) = \mathscr{B}(\mathbb{R})$. This extension is clearly a Lebesgue-Stieltjes measure because $\mu((a,b]) = F(b) - F(a)$ and F takes finite values.□

With this theorem, it is easy to construct the Lebesgue measure on $\mathscr{B}(\mathbb{R})$: simply take $F(x) = x$; then, the corresponding Lebesgue-Stieltjes measure, m, is the unique measure on $\mathscr{B}(\mathbb{R})$ assigning its length to each interval.

We can extend this measure to $\mathbb{R}^n$ with the theory developed so far: note that, by Lemma 4.2.8, $\mathscr{B}(\mathbb{R}^n) = (\mathscr{B}(\mathbb{R}))^n$. Now take each $\mu_k = m$ on Corollary 4.1.8 to obtain the unique measure on $\mathscr{B}(\mathbb{R}^n)$ assigning its volume to each rectangle. Finally, the Lebesgue measure on $\mathbb{R}^n$ is defined as the completion of this measure, in the sense of Definition 2.3.6.

This measure allows us to revisit Real Analysis from a measure-theoretic standpoint. So much so, that (proper) Riemann integration can be regarded as a particular case or Lebesgue integration, in the sense of the following theorem:

Theorem 5.1.4. *Let f be a bounded, real-valued function defined on a closed rectangle $R \subseteq \mathbb{R}^n$. Then,*

(i) f is Riemann integrable on R if, and only if, it is continuous almost everywhere on R with respect to the Lebesgue measure on $\mathbb{R}^n$.

(ii) If f is Riemann integrable on R, then it is integrable on R with respect to the Lebesgue measure on $\mathbb{R}^n$ and the two integrals are equal.

This result will be of little interest in this work, and its proof is not included due to space limitations. The interested reader can consult Section 1.7 of [6].

Starting here, the remaining part of the text has no longer been based on [6] and is mostly original instead.

More applications to Real Analysis are possible with this approach: Riemann-theoretic Fubini's Theorem is almost immediate from Corollary 4.1.9. Another classical theorem regarding integration is proved in Appendix C.

5.2 Existence of random variables and random samples

In Probability Theory, a **random variable** is defined as a (Borel) $\mathscr{F}$-measurable function $X\colon \Omega \to \mathbb{R}$, where $(\Omega, \mathscr{F}, P)$ is some probability space. The **distribution function** of X is then defined as the function $F_X\colon \mathbb{R} \to [0,1]$ given by $F_X(x) = P(\{X \le x\})$.

These two concepts are central in this theory, and the latter is always associated with the former. Usually, however, we are not interested in the random variable itself, but rather in its distribution function; in most introductory probability and statistics courses, random variables are introduced *via their distributions*, and not the other way around. Consider, for example, a normal distribution: we say that a given random variable X

follows a **normal distribution** with mean μ and variance σ^2 if

$$F_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{\mu-t}{\sigma}\right)^2} \, dt.$$

These random variables are widely used in mathematics, but it is not immediate that at least one of them exists: how do we know that there exist a probability space $(\Omega, \mathscr{F}, P)$ and random variable X on it such that F_X has the form specified above?

With the theory developed in this text, are able to construct any such function. Moreover, we will be able to guarantee that the probability space satisfies $\Omega = \mathbb{R}$ and $\mathscr{F} = \mathscr{B}(\mathbb{R})$.

Definition 5.2.1. *Consider a mapping $F\colon \mathbb{R} \to \mathbb{R}$. We will say that F is a* ***probabilistic distribution function*** *if F is a distribution function in the sense of Definition 5.1.1 that ranges between* 0 *and* 1*; namely, if*

(i) *F is an increasing function.*

(ii) *F is right-continuous; that is, if $x_n \downarrow x$, then $F(x_n) \to F(x)$ for every $x \in \mathbb{R}$.*

(iii) $\lim_{x\to-\infty} F(x) = 0$ *and* $\lim_{x\to+\infty} F(x) = 1$.

Proposition 5.2.2. *Let $F\colon \mathbb{R} \to \mathbb{R}$ be a probabilistic distribution function. Then, there exist a probability measure P on $\mathscr{B}(\mathbb{R})$ and a random variable $X\colon \mathbb{R} \to \mathbb{R}$ such that $F_X = F$.*

Proof. By Theorem 5.1.3, there exists a measure P on $\mathscr{B}(\mathbb{R})$ such that $F(x) = P((-\infty, x])$ for all $x \in \mathbb{R}$. To see that P is, in fact, a probability measure, take any sequence $x_n \uparrow +\infty$ ($x_n = n$ will do). Then,

$$P(\mathbb{R}) = \lim_n P((-\infty, x_n]) = \lim_n F(x_n) = 1.$$

Now simply define $X = id$. It is clear that X is measurable, and trivially

$$F_X(x) = P(\{X \le x\}) = P((-\infty, x]) = F(x). \qquad \square$$

A similar question arises when working with random samples, but the solution now is more complicated. Before proceeding, we remind the reader of some basic concepts related to this topic:

Let $\{X_t\}_{t\in T}$ be a family of random variables defined on the same probability space $(\Omega, \mathscr{F}, P)$. We say that this family is **independent** if, for any finite set of those random variables, $X_{t_1}, \ldots, X_{t_n}$ and any measurable sets $B_{t_1}, \ldots, B_{t_n}$, $B_{t_k} \in \mathscr{B}(\mathbb{R})$, we have

$$P\left(\left\{X_{t_1} \in B_{t_1}\right\} \cap \ldots \cap \left\{X_{t_n} \in B_{t_n}\right\}\right) = P\left(\left\{X_{t_1} \in B_{t_1}\right\}\right) \cdot \ldots \cdot P\left(\left\{X_{t_n} \in B_{t_n}\right\}\right),$$

where $\left\{X_{t_i} \in B_{t_i}\right\}$ denotes the (measurable) set $\left\{\omega \in \Omega \,\middle|\, X_{t_i}(\omega) \in B_{t_i}\right\} = X_{t_i}^{-1}(B_{t_i})$. We say that the family is **identically distributed** if $F_{X_t} = F_{X_l}$ for any two $t, l \in T$.

Following this notation, a **random sample** with a given distribution F is a countable, independent and identically distributed (with distribution F) family of random variables defined on the same probability space.

The concept of random sample is widely used in statistics, being at the core of many theorems and concepts. But, again, it is not trivial that there exists a random sample with a given distribution (for instance, a Bernoulli). The Kolmogorov Extension Theorem allows us to construct this, and even generalise it to the case where we have an arbitrary amount of random variables and each random variable has a different distribution.

Theorem 5.2.3. *Let T be an index set. Suppose that, for every given $t \in T$, we are given a probabilistic distribution function F_t. Then, there exists a probability space Ω and an independent family of random variables $\{X_t\}_{t\in T}$ (each defined on Ω) such that $F_{X_t} = F_t$.*

Proof. For every $t \in T$, let P_t be the Lebesgue-Stieltjes measure associated to F_t. It is clear that P_t is a probability measure on $\mathbb{R}$.
Identify $\Omega_t = \mathbb{R}$ and $\mathscr{F}_t = \mathscr{B}(\mathbb{R})$ for every t - so that each Ω_t is a separable, complete metric space. Also, for each finite $v \subseteq T$, write $v = \{t_1,\dots,t_n\}$, with $t_1 < \dots < t_n$. Then, $\Omega_v = \mathbb{R}^n$ and $\mathscr{F}_v = (\mathscr{B}(\mathbb{R}))^n$ for each finite $v \subseteq T$. Use Corollary 4.1.8 to obtain a probability measure P_v on $\mathscr{F}_v$ such that

$$P_v\left(A_{t_1}\times\cdots\times A_{t_n}\right) = P_{t_1}(A_{t_1})\cdot\ldots\cdot P_{t_n}(A_{t_n}),$$

for each finite family of sets $A_{t_k} \in \mathscr{F}_{t_k}$.
It is clear that the family of probability measures defined in this way is consistent: if $v \subseteq w$, we can suppose, for simplicity, that $w = v \cup \{t_{n+1}\}$, with $t_{n+1} > t_n$ (the "complete" result is proved very similarly, but with more cumbersome notation). Define the probability measure P' on $\mathscr{F}_v$ as $P'(B) = P_w(B^w)$. Note that $B^w = B \times \Omega_{t_{n+1 say,}}$. Then, for every measurable rectangle $B = A_{t_1}\times\cdots\times A_{t_n}$, we have

$$P'(B) = P_w(A_{t_1}\times\cdots\times A_{t_n}\times\Omega_{t_{n+1}}) = P_{t_1}(A_{t_1})\cdot\ldots\cdot P_{t_n}(A_{t_n})\cdot P_{t_{n+1}}(\Omega_{t_{n+1}}) = P_v(B).$$

Since P' and P_v agree on measurable rectangles, by the uniqueness part of Corollary 4.1.8 we have $P' = P_v$. Therefore, $P_v(B) = P_w(B^w)$ for every $B \in \mathscr{F}_v$.
Now use the Kolmogorov Extension Theorem to construct the hoped-for probability space $(\Omega,\mathscr{F},P)$, where $\Omega = \prod_t \Omega_t$ and $\mathscr{F} = \prod_t \mathscr{F}_t$.
For each fixed $t \in T$, define $X_t(\omega) = \omega_{\{t\}}$. It is clear that X_t is a measurable function, since $X_t^{-1}(B) = B_{\{t\}}$ is a measurable cylinder for each $B \in \mathscr{B}(\mathbb{R}) = \mathscr{F}_t$. Moreover, $F_{X_t} = F_t$, because

$$P(\{X_t \le x\}) = P\left((-\infty,x]_{\{t\}}\right) = P_{\{t\}}((-\infty,x]) = P_t((-\infty,x]) = F_t(x).$$

Finally, the family of random variables $\{X_t\}_{t\in T}$ is independent: consider finitely many random variables $X_{t_1},\dots,X_{t_n}$, and suppose they are ordered ($t_1 < \dots < t_n$). Write $v = \{t_1,\dots,t_n\}$. Let $B_1,\dots,B_n \in \mathscr{B}(\mathbb{R})$. Now note that, for each $k = 1,\dots,n$, we can regard $B_k \in \mathscr{B}(\mathbb{R}) = \mathscr{F}_{t_k} = \mathscr{F}_{\{t_k\}}$. Therefore, we can consider its retraction to Ω_v:

$$(B_k)^v = \Omega_{t_1}\times\cdots\times\Omega_{t_{k-1}}\times B_k\times\Omega_{t_{k+1}}\times\cdots\times\Omega_{t_n}.$$

Now write, for each k, $A_k = \left\{X_{t_k} \in B_k\right\} = (B_k)_{\{t_k\}} \in \mathscr{F}$. Note that $P(A_k) = P_{\{t_k\}}(B_k) = P_{t_k}(B_k)$. We can also regard A_k as having a higher base, common for all k:

$$A_k = (B_k)_{\{t_k\}} = \left((B_k)^v\right)_v = \left(\Omega_{t_1}\times\cdots\times\Omega_{t_{k-1}}\times B_k\times\Omega_{t_{k+1}}\times\cdots\times\Omega_{t_n}\right)_v.$$

It follows from Remark 4.2.2 that $A_1 \cap \cdots \cap A_n = ((B_1)^v \cap \cdots \cap (B_n)^v)_v = (B_1 \times \cdots \times B_n)_v$. Therefore,

$$P(A_1 \cap \cdots \cap A_n) = P_v(B_1 \times \cdots \times B_n) = P_{t_1}(B_1) \cdot \ldots \cdot P_{t_n}(B_n) = P(A_1) \cdot \ldots \cdot P(A_n).$$

Recall that each $A_k = \{X_{t_k} \in B_k\}$, so this is the condition of independency. □

Corollary 5.2.4. *Let F be a probabilistic distribution function. Then, there exists a random sample of F.*

Proof. In Theorem 5.2.3, take $T = \mathbb{Z}^+$ and $F_n = F$ for each $n \in \mathbb{Z}^+$. □

Conclusions

In this work, we have developed many tools in measure theory which have allowed us to obtain results in real analysis and probability theory. In Chapter 2, we introduced most of the necessary concepts regarding measure theory and integration. In Chapter 3, we develped strong tools relating this theory to function spaces and topology. In Chapter 4, we studied measures in higher dimensions, and this led us to the Kolmogorov Extension Theorem. Finally, in Chapter 5, we used the theory developed so far to gain some insight into real analysis and probability theory.

APPENDIX

A

LEMMAS ON SWITCHING LIMITS

Lemma A.1.1. *Let a_{nm} be a sequence of real numbers that is increasing with respect to both indices; that is, $n \geq n'$ implies $\forall m\colon a_{nm} \geq a_{n'm}$, and $m \geq m'$ implies $\forall n\colon a_{nm} \geq a_{nm'}$. Then,*

$$\lim_n \lim_m a_{nm} = \lim_m \lim_n a_{nm}$$

Proof. Since the sequence is increasing with respect to both indices, the result is equivalent to

$$\sup_n \sup_m a_{nm} = \sup_m \sup_n a_{nm}$$

We will first show that $\sup_n \sup_m a_{nm} = \sup_{n,m} a_{nm}$. For every n, define $k_n = \sup_m a_{nm}$. Also set $k = \sup_{n,m} a_{nm}$. It is clear that $\forall n\colon k_n \leq k$, and thus $\sup_n k_n \leq k$. On the other side, $\forall n \forall m\colon k_n \geq a_{nm}$. Therefore, taking suprema with respect to n, m, we have $\sup_{n,m} k_n \geq k$. But since k_n does not depend on m, $\sup_{n,m} k_n = \sup_n k_n$. Thus, $\sup_n \sup_m a_{nm} = \sup_{n,m} a_{nm}$.

Define $b_{nm} = a_{mn}$. The sequence b_{nm} is also increasing with respect to both indices. Hence,

$$\sup_n \sup_m a_{nm} = \sup_{n,m} a_{nm} = \sup_{n,m} b_{nm} = \sup_n \sup_m a_{mn} = \sup_m \sup_n a_{nm}$$ □

Corollary A.1.2. *Let a_{ij} be a sequence of nonnegative real numbers, where $i, j \in \mathbb{Z}^+$. Then,*

$$\sum_i \sum_j a_{ij} = \sum_j \sum_i a_{ij},$$

whether the expression is finite or not.

Proof. Define $S_{nm} = \sum_{i=1}^n \sum_{j=1}^m a_{ij}$. Then, the expression $\sum_i \sum_j a_{ij}$ corresponds to $\lim_n \lim_m S_{nm}$, and the expression $\sum_j \sum_i a_{ij}$ corresponds to $\lim_m \lim_n S_{nm}$. Since $\forall i \forall j\colon a_{ij} \geq 0$, the sequence S_{nm} is increasing with respect to both indices. Applying Lemma A.1.1 yields the desired result. □

APPENDIX B

RELATIONS BEWTEEN MEASURES

In Chapter 2 we studied measures mostly *intrinsically*: we constructed measure spaces and studied the properties of functions on fixed measure spaces. However, we did not study measures *extrinsically*, that is, how different measure spaces relate to each other. This is one of the main goals of this chapter.

B.1 Jordan-Hahn Decomposition Theorem

In this section we will develop a way of systematically studying σ-additive set functions defined on σ-fields. This class of functions is interesting because it contains all set functions of the kind $\lambda(A) = \int_A f \, d\mu$ for some Borel measurable function f such that $\int_\Omega f \, d\mu$ exists. Every such function can be written as the difference of two measures $\lambda = \lambda^+ - \lambda^-$: simply take $\lambda^+(A) = \int_A f^+ \, d\mu$ and $\lambda^-(A) = \int_A f^- \, d\mu$.

The Jordan-Hahn Decomposition Theorem states that the same is true for the entire class: any σ-additive set function on a σ-field can be expressed as the difference of two measures. Before stating it, we need some to develop one result that is very useful by itself, and will allow us to prove the main theorem:

Theorem B.1.1. *Let λ be a countably additive extended real-valued set function defined on a σ-field $\mathscr{F}$ of subsets of Ω. Then, λ assumes its maximum and minimum, that is, there exist $C, D \in \mathscr{F}$ such that*

$$\lambda(C) = \sup\{\lambda(A) : A \in \mathscr{F}\} \text{ and } \lambda(D) = \inf\{\lambda(A) : A \in \mathscr{F}\}$$

Proof. We will first consider the supremum. We may assume that $\lambda(A) < +\infty$ for every $A \in \mathscr{F}$, for if $\lambda(A_0) = \infty$, we take $C = A_0$. Take a sequence of sets $A_n \in \mathscr{F}$ such that $\lambda(A_n) \to \sup \lambda$, and set $A = \bigcup_n A_n$.

The idea of the proof is to "approximate" A by finitely many subsets $A_1, \ldots, A_n$ and take only the positive bits created by those subsets. More concretely, consider some fixed $n \in \mathbb{Z}^+$ and note that, for each $k \leq n$ and for each $a \in A$, either $a \in A_k$ or $a \in A_k^c$. Thus, we may write $A = \bigcup_{m=1}^{2^n} A_{nm}$, where the sets A_{nm} comprise all the different 2^n sets of the form $A_1^* \cap \cdots \cap A_n^*$, where A_k^* is either A_k or A_k^c. From this construction, we can observe that:

1. The sets A_{nm} are disjoint.
2. For every $k \leq n$, A_k is the finite union of some of the sets A_{nm}.
3. If $n' > n$, each A_{nm} is a subset of some $A_{n'm'}$.

Define B_n as the union of all sets A_{nm} such that $\lambda(A_{nm}) \geq 0$. From 2, it follows that $\lambda(A_n) \leq \lambda(B_n)$ and from 3 we have that, for $k > n$, $\bigcup_{k=n}^{r} B_k$ can be written as the disjoint union of B_n and some sets $A_{n'm'}$. Thus, $\lambda(\bigcup_{k=n}^{r} B_k) \geq \lambda(B_n)$. Therefore, we have

$$\sup \lambda = \lim_n \lambda(A_n) \leq \lim_n \lambda(B_n) \leq \lambda\left(\bigcup_{k=n}^{r} B_n\right) \uparrow_r \lambda\left(\bigcup_{k=n}^{+\infty} B_k\right).$$

If we set $C = \limsup_n B_n$, since $\lambda\left(\bigcup_{k=n}^{+\infty} B_k\right) \downarrow \lambda(C)$, we have $\sup \lambda \leq \lambda(C) \leq \sup \lambda$.
The set D is obtained by applying the result proved so far to $-\lambda$. □

We now have all tools required to develop our theorem. Without further delay:

Theorem B.1.2 (Jordan-Hahn Decomposition Theorem). *Let λ be a countably additive set function defined on a σ-field $\mathscr{F}$. Define the set functions*

$$\lambda^+(B) = \sup\{\lambda(A) : A \subseteq B, A \in \mathscr{F}\} \text{ and } \lambda^-(B) = -\inf\{\lambda(A) : A \subseteq B, A \in \mathscr{F}\}.$$

Then, λ^+ and λ^- are measures on $\mathscr{F}$ and $\lambda = \lambda^+ - \lambda^-$.

Proof. First, suppose, without loss of generality, that $\lambda < +\infty$ (by definition of countably additive set function, $+\infty$ and $-\infty$ cannot both be in the range of λ, and if $\lambda > -\infty$, we can apply the result to $-\lambda$). Let $C \in \mathscr{F}$ be a set such that, for all $A \in \mathscr{F}$,

$$\lambda(A \cap C) \geq 0 \text{ and } \lambda(A \cap C^c) \leq 0.$$

Such a set exists because of Theorem B.1.1: take $C \in \mathscr{F}$ satisfying $\lambda(C) = \sup \lambda$. Then, if $\lambda(A \cap C) < 0$, we have $\sup \lambda = \lambda(C) = \lambda(C \setminus (C \cap A)) + \lambda(A \cap C) < \lambda(C \setminus (C \cap A))$; and if $\lambda(A \cap C^c) > 0$, then $\lambda(C \cup (A \cap C^c)) = \lambda(C) + \lambda(A \cap C^c) > \lambda(C) = \sup \lambda$.
From this, we will see that $\lambda^+(A) = \lambda(A \cap C)$ and $-\lambda^-(A) = \lambda(A \cap C^c)$:
Inequalities $\lambda(A \cap C) \leq \lambda^+(A), \lambda(A \cap C^c) \geq \lambda^-(A)$ are immediate from the definitions of λ^+ and λ^-. To see the equality, note that, if $B \subseteq A, B \in \mathscr{F}$,

$$\lambda(B) = \lambda(B \cap C) + \lambda(B \cap C^c) \leq \lambda(B \cap C) \leq \lambda(B \cap C) + \lambda((A \setminus B) \cap C) = \lambda(A \cap C).$$

Thus, $\lambda^+(A) \leq \lambda(A \cap C)$. Similarly, $-\lambda^-(A) \geq \lambda(A \cap C^c)$.
From here, it clearly follows that both λ^+ and λ^- are measures and that $\lambda = \lambda^+ - \lambda^-$. □

Before finishing the section, we can extract a few additional results.

Corollary B.1.3. *Let λ be a countably additive extended real-valued set function on the σ-field $\mathscr{F}$. Then,*

(i) λ is the difference of two measures, at least one of which is finite.

(ii) If λ is finite, then it is bounded.

(iii) There is a set $C \in \mathscr{F}$ such that $\lambda(A \cap C) \geq 0$ and $\lambda(A \cap C^c) \leq 0$ for all $A \in \mathscr{F}$.

(iv) If $E \in \mathscr{F}$ is another set satisfying that $\lambda(A \cap E) \geq 0$ and $\lambda(A \cap E^c) \leq 0$ for all $A \in \mathscr{F}$, then $\lambda^+(A) = \lambda(A \cap E)$ and $\lambda^-(A) = \lambda(A \cap E^c)$ for all $A \in \mathscr{F}$.

(v) If E is one such set, then $|\lambda|(C \triangle E) = 0$, where $|\lambda| = \lambda^+ + \lambda^-$.

Proof. 1. If $\lambda < +\infty$, then $\lambda^+ < +\infty$, and if $\lambda > -\infty$, then $\lambda^- < +\infty$.

2. Consequence of the previous section and the fact that every finite measure is bounded.

3. Consequence of Theorem B.1.1, as shown in the proof of Theorem B.1.2.

4. The set C in the proof of B.1.2 was imposed no other hypotheses except that $C \in \mathscr{F}$ and $\lambda(A \cap C) \geq 0, \lambda(A \cap C^c) \leq 0$. Thus, any such set E will satisfy $\lambda^+(A) = \lambda(A \cap E), \lambda^-(A) = \lambda(A \cap E^c)$.

5. First note that by the property satisfied by C and E, we have $0 \leq \lambda(C \cap E^c) \leq 0$, and thus, by the previous section we have $\lambda^+(E^c) = \lambda^-(C) = 0$. Therefore, $0 \leq \lambda^+(C \cap E^c) \leq \lambda^+(E^c) = 0$, and $0 \leq \lambda^-(C \cap E^c) \leq \lambda^-(C) = 0$. □

The Jordan-Hahn Decomposition Theorem motivates the use of the expression **signed measure** for denoting any countably additive extended-real valued set function defined on a σ-field.

Given a signed measure λ, we call λ^+ its **upper variation**, λ^- its **lower variation**, and $|\lambda| = \lambda^+ + \lambda^-$ its **total variation**.

Remark B.1.4. *Note that, for every $A \in \mathscr{F}$, $|\lambda(A)| \leq |\lambda|(A)$:*

$$|\lambda(A)| = \left|\lambda^+(A) - \lambda^-(A)\right| \leq \lambda^+(A) + \lambda^-(A) = |\lambda|(A).$$

Additionally, $|\lambda|(A) = 0$ if, and only if, $\lambda(B) = 0$ for every $B \subseteq A, B \in \mathscr{F}$.

Remark B.1.5. *For every signed measure λ and every $\alpha \in \overline{\mathbb{R}}$, one has $|\alpha\lambda| = |\alpha||\lambda|$. Additionally, if τ is another signed measure such that $\lambda + \tau$ is well defined, then $|\lambda + \tau|(A) \leq |\lambda|(A) + |\tau|(A)$ for every measurable A.*

B.2 Absolute continuity and singularity of signed measures

We are now ready to study relations between measures. From this study will arise two important results: the Radon-Nikodým Theorem and the Lebesgue Decomposition Theorem.
In this section, we are to introduce two important concepts, that can be regarded, in some sense, as opposite. A (signed) measure is said to be *absolutely continuous* with respect to another signed measure whenever it can have no effect on sets that are null with respect to the former. We might also be interested on (signed) measures that *only* have effect on null sets with respect to the former, and these are called *singular*.

Definition B.2.1. *Let $\mathscr{F}$ be a σ-field, μ be a measure on $\mathscr{F}$ and ν be a signed measure on $\mathscr{F}$. We say that ν is* ***absolutely continuous*** *with respect to μ if $\mu(A) = 0$ implies $\nu(A) = 0$, and we denote it by $\nu \ll \mu$. If λ_1 and λ_2 are signed measures, we say that λ_1 is* ***absolutely continuous*** *with respect to λ_2 whenever $\lambda_1 \ll |\lambda_2|$.*
We say that ν is ***singular*** *with respect to μ if there exists some $A \in \mathscr{F}$ such that $\mu(A) = 0$ and $\nu(A^c) = 0$, and we denote it by $\nu \perp \mu$. If λ_1 and λ_2 are signed measures, we say that λ_1 is* ***singular*** *with respect to λ_2 whenever $|\lambda_1| \perp |\lambda_2|$.*

It is clear that if $\lambda_1 \perp \lambda_2$, then $\lambda_2 \perp \lambda_1$. As expected, there are relations between the two concepts. We capture this in the following result:

Lemma B.2.2. *Let λ_1, λ_2 and ν be signed measures defined on a σ-field $\mathscr{F}$.*

(i) If $\lambda_1 \perp \nu$ and $\lambda_2 \perp \nu$, then, for every $\alpha_1, \alpha_2 \in \overline{\mathbb{R}}$ such that $\alpha_1\lambda_1 + \alpha_2\lambda_2$ is well-defined, we have that $\alpha\lambda_1 + \beta\lambda_2 \perp \nu$ too.

(ii) $\lambda_1 \ll \nu$ if, and only if, $|\lambda_1| \ll \nu$.

(iii) If $\lambda_1 \ll \nu$ and $\lambda_2 \perp \nu$, then $\lambda_1 \perp \lambda_2$.

(iv) If $\lambda_1 \ll \nu$ and $\lambda_1 \perp \nu$, then $\lambda_1 \equiv 0$.

(v) If λ_1 is finite, then $\lambda_1 \ll \nu$ if, and only if, $\lim_{|\nu|(A)\to 0} \lambda_1(A) = 0$.

Proof.

1. Let A_1 and A_2 be sets in $\mathscr{F}$ such that $|\lambda_1|(A_1) = |\lambda_2|(A_2) = 0$ and $|\nu|(A_1^c) = |\nu|(A_2^c) = 0$. Let $B = A_1 \cap A_2$. It is clear that $|\nu|(B^c) = 0$ and $|\lambda_1|(B) = |\lambda_2|(B) = 0$. Then, by Remark B.1.5, we have $|\alpha_1\lambda_1 + \alpha_2\lambda_2|(B) \le |\alpha_1\lambda_1|(B) + |\alpha_2\lambda_2|(B) = |\alpha_1| \cdot 0 + |\alpha_2| \cdot 0 = 0$.

2. Immediate by Remark B.1.4.

3. Let $A \in \mathscr{F}$ be a set such that $|\lambda_2|(A) = |\nu|(A^c) = 0$. Since $|\lambda_1| \ll |\nu|$ (by B.2.2.(ii)), it must be $|\lambda_1|(A^c) = 0$. Thus, $\lambda_1 \perp \lambda_2$.

4. By B.2.2.(iii), we have $\lambda_1 \perp \lambda_1$. It follows that there exists some $A \in \mathscr{F}$ such that $|\lambda_1|(\Omega) = |\lambda_1|(A) + |\lambda_1|(A^c) = 0 + 0 = 0$. Thus, $|\lambda_1| \equiv 0$, whence $\lambda_1 \equiv 0$

5. Suppose that $\lim_{\nu(A)\to 0} \lambda_1(A) = 0$. If $B \in \mathscr{F}$ is a set such that $|\nu|(B) = 0$, it follows that $|\lambda_1(B)| < \varepsilon$ for all $\varepsilon > 0$, whence $\lambda_1(B) = 0$. Conversely, suppose that there exists some $\varepsilon > 0$ such that, for every $n \in \mathbb{N}$ there exists some $A_n \in \mathscr{F}$ with

$|\nu|(A_n) < 2^{-n}$ and $|\lambda_1(A_n)| \geq \varepsilon$. Note that $|\lambda_1|(A_n) \geq |\lambda(A_n)| \geq \varepsilon$. Thus, if we define $A = \limsup_n A_n$, we have $|\lambda_1|(\bigcup_{k\geq n} A_k) \geq |\lambda_1|(A_n) \geq \varepsilon$, whence

$$|\lambda_1|(A) = \lim_{n\to+\infty} \left(\bigcup_{k\geq n} A_k\right) \geq \varepsilon.$$

However, since $\sum_n |\nu|(A_n) < +\infty$, by the Borel-Cantelli Lemma, it must be $|\nu|(A) = 0$, a contradiction. □

Note that if λ can be written as the integral of some Borel measurable function g, that is, $\lambda(A) = \int_A g\,d\mu$ for every $A \in \mathscr{F}$, then $\lambda \ll \mu$. The Radon-Nikodým Theorem states the converse result, under the hypothesis that μ be σ-finite:

Theorem B.2.3 (Radon-Nikodým Theorem). *Let μ be a σ-finite measure defined on a measurable space $(\Omega, \mathscr{F})$. Let λ be a signed measure that is absolutely continuous with respect to μ. Then, there exists a Borel measurable function $g\colon \Omega \to \overline{\mathbb{R}}$ such that*

$$\lambda(A) = \int_A g\,d\mu \quad \textit{for every } A \in \mathscr{F}.$$

If h is another such function, then $g = h$ μ-a.e.

Proof. Uniqueness is a direct consequence of Corollary 2.4.29. We will start the proof with strict hypotheses to μ and λ and work upwards.

1. Suppose λ and μ are finite measures.

   This theorem states the existence of a function satisfying certain “abstract” properties. Many such theorems can be proved by using Zorn’s lemma, and that is what we are going to do. Firstly, we need to construct the set and a partial order in it whose maximal element is to be our candidate function. With this in mind, consider the set

   $$G = \left\{ g\colon \Omega \to \overline{\mathbb{R}} \,\middle|\, g \text{ is Borel measurable}, g \geq 0 \text{ and } \lambda(A) \geq \int_A g\,d\mu \text{ for every } A \in \mathscr{F} \right\},$$

   and $\mathscr{G} = G/\sim$, where $\sim$ is the usual equivalence relation identifying functions that coincide μ-a.e. The choice for G makes sense because it transforms our desired property into a more easily “partially orderable” one, and we need to take the quotient because Theorem 2.4.28 only ensures inequalities μ-a.e. The set constructed is nonempty because $0 \in \mathscr{G}$ (this is the reason why we choose $\geq$ instead of $\leq$ in the definition of G). Now we need to partially order our set. Keeping in mind that we want our maximal element g to be our candidate, and that we have ensured that $\lambda(A) \geq \int_A g\,d\mu$, we would like the equality not to be strict. We want, then, integrals of g to be “the biggest” as possible. Thus, a good candidate for partial ordering $\leq^*$ would be $h_1 \leq^* h_2$ whenever $\int_A h_1\,d\mu \leq \int_A h_2\,d\mu$ for all $A \in \mathscr{F}$. However, by Theorem 2.4.28, this is equivalent to simply $h_1 \leq h_2$ μ-a.e. in the standard sense. Therefore we will define our ordering in $\mathscr{G}$ as $h_1 \leq^* h_2$ whenever $h_1 \leq h_2$ μ-almost everywhere on Ω.

   Having a candidate partially ordered set, we will now use Zorn’s lemma to see it has a maximal element. Let $\mathscr{I}$ be a chain of $\mathscr{G}$. Let $M = \sup_{f\in\mathscr{I}} \left\{\int_\Omega f\,d\mu\right\}$ and

$f_1, f_2, \ldots$ a sequence of functions in $\mathscr{I}$ such that $\int_\Omega f_n d\mu \uparrow M$. Since $\mathscr{I}$ is a chain, the sequence of functions is necessarily increasing μ-a.e. Thus, we can define $f = \lim_n f_n$ almost everywhere and $f = 0$ on the set where the sequence is not monotone. By the Extended Monotone Convergence Theorem, $\int_\Omega f d\mu = M$. This function is an upper bound of $\mathscr{I}$: Let $g \in \mathscr{I}$. If $g \geq^* f_n$ for all n, then $g \geq^* f$, and thus $\int_\Omega g d\mu = M = \int_\Omega f d\mu$, but then $g = f$ μ-a.e. If $g <^* f_n$, since $f_n \uparrow f$ μ-a.e., then $g <^* f$.

By Zorn's lemma, $\mathscr{G}$ has a maximal element, g. We will now show that this is the function we were looking for. Define a measure on $\mathscr{F}$ by $\nu(A) = \lambda(A) - \int_A g d\mu$. Since λ is absolutely continuous with respect to μ, so is ν. Suppose, by way of contradiction, that $\nu(\Omega) > 0$, and set $k = 2\mu(\Omega)/\nu(\Omega) > 0$, so that

$$k\nu(\Omega) - \mu(\Omega) > 0$$

Define a signed measure $\eta = k\nu - \mu$, which is absolutely continuous with respect to μ, and use Corollary B.1.3.(iii) to obtain a set $D \in \mathscr{F}$ such that $\eta(A \cap D) \geq 0$ for all $A \in \mathscr{F}$. By Corollary B.1.3.(v), we have that $\eta(D) = \eta(\Omega \cap D) = \eta^+(\Omega) \geq \eta(\Omega) > 0$. Since η is absolutely continuous with respect to μ, then $\mu(D) > 0$ (if $\mu(D)$ was 0, so would be $\eta(D)$).

Finally, define the function $h = g + \frac{1}{k} I_D$. Note that h is Borel measurable, nonnegative and since $\mu(D) > 0$, $h >^* g$. Now, for any $A \in \mathscr{F}$, we have

$$\lambda(A) - \int_A h d\mu = \nu(A) - \frac{1}{k}\mu(A \cap D) \geq \frac{1}{k}\left(k\nu(A \cap D) - \mu(A \cap D)\right) = \frac{1}{k}\eta(A \cap D) \geq 0.$$

Therefore, $h \in G$, contradicting the maximality of g. It must be that $\nu(\Omega) = 0$, and thus $\nu(A) = \lambda(A) - \int_A g d\mu = 0$ for every $A \in \mathscr{F}$.

2. Suppose λ is a σ-finite measure and μ is a finite measure.

   Decompose λ into countably many finite measures: $\lambda = \sum_n \lambda_n$. For every n, λ_n is absolutely continuous with respect to μ. Apply 1 to obtain a function g_n that is nonnegative μ-a.e. Then, $g = \sum_n g_n$ is the desired function: for every $A \in \mathscr{F}$,

   $$\lambda(A) = \sum_n \lambda_n(A) = \sum_n \int_A g_n d\mu = \int_A g d\mu,$$

   where in the last step we used 2.4.21.(i).

3. Suppose λ is an arbitrary measure and μ is a finite measure.

   Given some $C \in \mathscr{F}$, define the σ-field $\mathscr{F}_C = \{B \cap C \colon B \in \mathscr{F}\}$. Define the measures λ_C and μ_C over $\mathscr{F}_C$ as $\lambda_C = \lambda|_{F_C}$, $\mu_C = \mu|_{F_C}$. Note that λ_C is absolutely continuous with respect to μ_C. Let $\mathscr{S}$ be the class of sets $C \in \mathscr{F}$ such that λ_C is σ-finite. $\mathscr{S}$ is not empty since $\emptyset \in \mathscr{S}$. Note that $\mathscr{S}$ is closed under countable unions: if $S_1, S_2, \ldots$ is a sequence of sets in $\mathscr{S}$ and $S = \bigcup_n S_n$, we can write $S_n = \bigcup_m A_{nm}$, where $\lambda(A_{nm}) < \infty$. Thus, $S = \bigcup_{n,m} A_{nm}$, showing that $S \in \mathscr{S}$.

   Let $M = \sup\{\mu(A) \colon A \in \mathscr{S}\}$, and consider a sequence of sets $B_1, B_2, \ldots \in \mathscr{S}$ such that $\mu(S_n) \uparrow M$. Let $B = \bigcup_n S_n \in \mathscr{S}$. Then, for all n, $M \geq \mu(B) \geq \mu(S_n)$, whence $\mu(B) = M$.

Apply 2 to λ_B and μ_B to obtain a function g'. Define $g = g' + \infty I_{B^c}$. Now, for any set $A \in \mathscr{F}$,

$$\lambda(A) = \lambda(A \cap B) + \lambda(A \setminus B) = \int_{A\cap B} g' d\mu + \lambda(A \setminus B) = \int_{A\cap B} g d\mu + \lambda(A \cap B).$$

We need only to show that $\lambda(A \setminus B) = \int_{A\setminus B} g d\mu$. If $\mu(A \setminus B) = 0$, both values are clearly 0. If $\mu(A \setminus B) > 0$, clearly $\int_{A\setminus B} g d\mu = +\infty$, and it follows too that $\lambda(A \setminus B) = +\infty$: suppose $\lambda(A \setminus B) < \infty$. Then $A \cup B = B \cup (A \setminus B) \in \mathscr{S}$, because we can decompose B into countably many sets $B_1, B_2, \dots$ such that $\lambda(B_n) < \infty$, and thus $A \cup B = \bigcup_n B_n \cup (A \setminus B)$, which is a countable union. However, $M \ge \mu(A \cup B) = \mu(B) + \mu(A \setminus B) > \mu(B) = M$, a contradiction.

4. Suppose λ is an arbitrary measure and μ is a σ-finite measure.

   Since μ is σ-finite, Ω can be decomposed into countably many disjoint sets $A_1, A_2, \dots$ such that $\mu(A_n) < \infty$ for every n. Define $\lambda_n = \lambda_{A_n}$, $\mu_n = \mu_{A_n}$ (with the notation established in 3). It is clear that λ_n is σ-finite with respect to μ_n for every n. Apply 3 to obtain a function g_n that is nonnegative μ-a.e. Then, $g = \sum_n g_n I_{A_n}$ is the desired function, as we will show:

   First, consider any $A \in \mathscr{F}$, and

$$\int_A g_n I_{A_n} d\mu_n = \int_{A_n \cap A} g_n d\mu_n.$$

   Note that, if we regard $(A_n \cap A, \mathscr{F}_{A_n\cap A})$ as a subspace of Ω (as in Remark 2.4.16), and write $\mathscr{F}_n = \mathscr{F}_{A_n\cap A} = \{B \cap (A_n \cap A) \colon B \in \mathscr{F}\}$, then $\mu|_{\mathscr{F}_n} \equiv \mu_n|_{\mathscr{F}_n}$. Therefore, their integrals coincide; that is, $\int_{A_n\cap A} g_n d\mu_n = \int_{A_n\cap A} g_n d\mu$, and this last integral is simply $\int_A g_n I_{A_n} d\mu$.

   It now follows that, by Corollary 2.4.21.(i),

$$\lambda(A) = \sum_n \lambda_n(A) = \sum_n \int_A g_n I_n d\mu = \int_A g d\mu$$

5. Suppose λ is an arbitrary measure and μ is a σ-finite measure.

   Use Jordan-Hahn Decomposition Theorem to split λ as the difference of two measures: $\lambda = \lambda^+ - \lambda^-$. Note that $|\lambda|$ is absolutely continuous with respect to μ by Remark B.1.4 and, hence, so are λ^+ and λ^-. Apply 4 to λ^+ and λ^-, to obtain g^+ and g^-, respectively. Since at least one of λ^+ and λ^- is finite, so is at least one of $\int_\Omega g^+ \, d\mu$ or $\int_\Omega g^- \, d\mu$. Thus, if we define the function $g = g^+ - g^-$, its integral $\int_\Omega g \, d\mu$ exists. Therefore, for every $A \in \mathscr{F}$, $\int_A g d\mu$ exists and

$$\lambda(A) = \lambda^+(A) - \lambda^-(A) = \int_A g^+ \, d\mu - \int_A g^- \, d\mu = \int_A g \, d\mu. \qquad \square$$

A different approach to proving the Radon-Nikodym Theorem is followed in [9].

Finally, we are ready to give a second decomposition theorem, which relates the two concepts introduced in this section.

Theorem B.2.4 (Lebesgue Decomposition Theorem). *Let $\mathscr{F}$ be a σ-field and μ a σ-finite measure on $\mathscr{F}$. Let λ be a signed measure on $\mathscr{F}$ such that $|\lambda|$ is σ-finite. Then, there exists a unique decomposition $\lambda = \lambda_1 + \lambda_2$, where λ_1 and λ_2 are signed measures such that $\lambda_1 \ll \mu$ and $\lambda_2 \perp \mu$.*
Additionally, if λ is a measure, then λ_1 and λ_2 are measures.

Proof. First, suppose that λ is a measure. Define $m = \mu + \lambda$, which is a σ-finite measure (if μ is finite on $A_1, A_2, \ldots$ and λ is finite on $B_1, B_2, \ldots$, then m is finite on the sets $C_{nm} = A_n \cap B_m$, which there are countably many of, and cover Ω). Additionally, both μ and λ are absolutely continuous with respect to m.
By the Radon-Nikodym Theorem, there exist a nonnegative, Borel measurable function f such that

$$\mu(A) = \int_A f \, dm.$$

Define the sets $B = \{\omega \in \Omega | f(\omega) > 0\}$ and $C = B^c = \{\omega \in \Omega | f(\omega) = 0\}$. Define the measures

$$\lambda_1(A) = \lambda(A \cap B), \quad \lambda_2(A) = \lambda(A \cap C).$$

It is clear that $\lambda = \lambda_1 + \lambda_2$. Additionally, if $\mu(A) = \int_A f \, dm = 0$, it must be that $f = 0$ m-a.e on A. Since $f > 0$ on $A \cap B$, it must be that $m(A \cap B) = 0$, whence $\lambda_1(A) = \lambda(A \cap B) = 0$. Thus, $\lambda_1 \ll \mu$. Finally , $\lambda_2(B) = \lambda(B \cap C) = \lambda(\emptyset) = 0$, and $\mu(B^c) = \mu(C) = \int_C f \, dm = \int_C 0 \, dm = 0$.
For the general case, use Theorem B.1.2 to split $\lambda = \lambda^+ - \lambda^-$. Since $|\lambda|$ is σ-finite, so are λ^+ and λ^-. Use the result proved so far to decompose $\lambda^+ = \lambda_1^+ + \lambda_2^+$, $\lambda^- = \lambda_1^- + \lambda_2^-$, with $\lambda_1^+, \lambda_1^- \ll \mu$ and $\lambda_2^+, \lambda_2^- \perp \mu$.
Since at least one of λ^+, λ^- is finite, we can define the signed measures $\lambda_1 = \lambda_1^+ - \lambda_1^-$ and $\lambda_2 = \lambda_2^+ - \lambda_2^-$ (the expression $+\infty - \infty$ is never attained), so that $\lambda = \lambda_1 + \lambda_2$. It is easy to check that $\lambda_1 \ll \mu$ and, by Lemma B.2.2.(i), $\lambda_2 \perp \mu$. $\square$

APPENDIX C

A Brief Application to Analysis

Proposition C.1.1. *Let $C \subseteq \mathbb{R}^{n-1}$ be a compact set (with the usual topology). Let $l, u\colon C \to \mathbb{R}$ be continuous functions such that $l(x) \le u(x)\ \forall x \in C$. Consider the set*

$$S = \{(x,t) \in \mathbb{R}^{n-1} \times \mathbb{R} \,|\, x \in C, l(x) \le t \le u(x)\}.$$

Let $f\colon S \to \mathbb{R}$ be a bounded, continuous function. Then, f is integrable on S and

$$\int_S f = \int_{x \in C} \int_{t=l(x)}^{t=u(x)} f(x,t),$$

where the integration is understood to be performed in the Riemann sense.

Proof. We know, from Theorem 5.1.4, that Riemann integration coincides with Lebesgue integration (with respect to the Lebesgue measure) on "classical" rectangles of $\mathbb{R}^n$ (that is, cartesian products of intervals). Riemann integration is usually defined on classical rectangles and then extended to bounded sets via indicator functions (see [10]). Thus, both kinds of integration will coincide on bounded sets.

It follows from Heine-Borel and Weierstrass Theorems that S is a bounded subset of $\mathbb{R}^n$ (Heine-Borel guarantees the boundedness of C, and Weierstrass guarantees the boundedness of limit functions $a(x), b(x)$). Thus, we can apply all techniques regarding Lebesgue integration.

Additionally, S is closed because of the following reasoning: if $g\colon X \to \mathbb{R}$ is a continuous function, then the epigraph $E(f) = \{(x,y) \in X \times \mathbb{R} \,\big|\, g(x) \le y\}$ is a closed set of the product topological space $X \times \mathbb{R}$. Similarly, so is the hypograph $H(f) = \{(x,y) \in X \times \mathbb{R} \,\big|\, g(x) \ge y\}$. Then, $S = E(l) \cap H(u)$ is closed. By the Heine-Borel theorem, S is compact.

Extend f to $\mathbb{R}^n$ setting its value to 0 outside of C. Now $I_S f$, regarded as a function from $\mathbb{R}^n$ to $\mathbb{R}$, is Borel measurable (since S is closed, it is Borel measurable) and bounded because f is (by Weierstrass Theorem). Since S is also bounded, $I_S f$ is integrable on $\mathbb{R}^n$.

Now split $f = f^+ - f^-$. Use Remark 4.1.6 to see that

$$\int_S f^+ = \int_{\mathbb{R}^{n-1}} \left(\int_{S^x} f^+(x,t)\ dt \right) dx.$$

Now note that $S^x = [l(x), u(x)]$ if $x \in C$ and $S^x = \emptyset$ otherwise. Therefore,

$$\int_S f^+ = \int_C \int_{l(x)}^{u(x)} f^+(x,t)\ dt\ dx.$$

A similar reasoning applied to f^- yields the desired result expanding and regrouping the expression $\int_S f = \int_S f^+ - \int_S f^-$. $\square$

Bibliography

[1] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, ser. Universitext. Springer New York, 2010. [Online]. Available: https://books.google.es/books?id=GAA2XqOIIGoC 1, 3.1

[2] W. Rudin, *Real and Complex Analysis*, ser. McGraw-Hill series in higher mathematics. Tata McGraw-Hill, 2006. [Online]. Available: https://books.google.es/books?id=3d8umJoRd08C 1

[3] F. W. Hammond, C. Sbert, and J. Duran, "Two nonlocal variational models for retinex image decomposition," in *International Conference on Computer Vision Theory and Applications.* SCITEPRESS, 2024. 1

[4] J. Duran, A. Buades, B. Coll, and C. Sbert, "A nonlocal variational model for pan-sharpening image fusion," *SIAM Journal on Imaging Sciences*, vol. 7, no. 2, pp. 761–796, 2014. 1

[5] J. von Plato, "Chapter 75 - a.n. kolmogorov, grundbegriffe der wahrscheinlichkeitsrechnung (1933)," in *Landmark Writings in Western Mathematics 1640-1940*, I. Grattan-Guinness, R. Cooke, L. Corry, P. Crépel, and N. Guicciardini, Eds. Amsterdam: Elsevier Science, 2005, pp. 960–969. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B978044450871350156X 1

[6] R. Ash, *Real Analysis and Probability*, ser. Bibliographie-P. Academic Press, 1972. [Online]. Available: https://books.google.es/books?id=0iTvAAAAMAAJ 1, 9, 10, 12, 14, 3, 4, 4.2, 8, 62, 5.1

[7] S. Axler, *Measure, Integration & Real Analysis*, ser. Graduate Texts in Mathematics. Springer International Publishing, 2019. [Online]. Available: https://books.google.es/books?id=DQ3BDwAAQBAJ 2.2

[8] H. Lebesgue, "Sur une généralisation de l'intégrale définie." *C. R. Acad. Sci., Paris*, vol. 132, pp. 1025–1028, 1901. 6

[9] A. Kolmogorov and S. Fomin, *Elements of the Theory of Functions and Functional Analysis*, ser. Elements of the Theory of Functions and Functional Analysis. Graylock Press, 1957, no. v. 1. [Online]. Available: https://books.google.es/books?id=oyJnbg6aTDIC B.2

[10] M. González Hidalgo, "Notes of the course "càlcul integral en diverses variables"," 2022. 74