Lebesgue Measure & Integration

Kexing Ying

January 11, 2021

Contents

1 Motivation

We recall from **Analysis I** the definition of the Darboux integral. While this notion of integration was sufficient for our use case last year, as we shall see, there are some limitations with this notion of integration. These limitations will be addressed by the means of measure theory.

Definition 1.1 (Darboux Integrable)**.** A function $f : [a, b] \to \mathbb{R}$ is called Darboux integrable if for any partition $\mathcal{P} = \{a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b\}$ for some $n \geq 1$ if $[a, b]$, by defining the lower and upper Darboux sums,

$$L(f, \mathcal{P}) = \sum_{i=1}^{n} (t_i - t_{i-1}) \inf_{t \in [t_{i-1}, t_i]} f(t),$$

and

$$U(f, \mathcal{P}) = \sum_{i=1}^{n} (t_i - t_{i-1}) \sup_{t \in [t_{i-1}, t_i]} f(t),$$

one has

$$\sup_{\mathcal{P}} L(f, \mathcal{P}) = \inf_{\mathcal{P}} U(f, \mathcal{P}).$$

If this is the case we define the integral of f over $[a, b]$ to be this value, i.e.

$$\int_a^b f := \sup_{\mathcal{P}} L(f, \mathcal{P}) = \inf_{\mathcal{P}} U(f, \mathcal{P}).$$

Many functions are Darboux integrable and in fact, as demonstrated last year, all functions in $C^{\circ}_{pw}([a, b])$, that is piecewise continuous functions on $[a, b]$ are Darboux integrable. Nonetheless, however, the class of Darboux integrable functions is also rather limited.

Consider the Dirichlet function

$$\mathbf{1}_{\mathbb{Q}}(x) := \begin{cases} 1, & x \in \mathbb{Q}; \\ 0, & x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

That is, the indicator function for $\mathbb{Q}$. We see that $\mathbf{1}_{\mathbb{Q}}$ is not Darboux integrable since both $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ are dense in $\mathbb{R}$ and so, for any partition $\mathcal{P}$ of $[a, b]$, $L(\mathbf{1}_{\mathbb{Q}}, \mathcal{P}) = 0$ while $U(\mathbf{1}_{\mathbb{Q}}, \mathcal{P}) = 1$. This is not ideal, since, as $\mathbb{Q}$ is countable while $\mathbb{R} \setminus \mathbb{Q}$ is not, we intuitively expect that a satisfactory theory of integration would assign $\int_a^b \mathbf{1}_{\mathbb{Q}} = 0$.

Moreover, by defining $P = (q_n)_{n \in \mathbb{N}} \subseteq \mathbb{Q}$ be some enumeration of $\mathbb{Q} \cap [a, b]$, we can define the following sequence of functions,

$$f_n(x) := \begin{cases} 1, & x \in \{q_0, \cdots, q_n\}; \\ 0, & \text{otherwise.} \end{cases}$$

It is not difficult to see that $\int_a^b f_n = 0$ for all n and $f_n \to \mathbf{1}_{\mathbb{Q}}$ pointwise. However, this implies

$$0 = \lim_{n \to \infty} \int_a^b f_n \neq \int_a^b \lim_{n \to \infty} f_n = \int_a^b \mathbf{1}_{\mathbb{Q}},$$

and in fact, the right hand side is not even defined (as $\mathbf{1}_{\mathbb{Q}}$ is not Darboux integrable)!

To solve this issue we will introduce the notion of the Lebesgue measure and furthermore, its associated Lebesgue integral which extends our Darboux integral such that it has the "nice" properties we desire.

We will in this course also look at L^p spaces. From the perspective of analysis, it is often convenient to work in Banach spaces (complete normed vector spaces) such that we can utilise many existing theorems we have proved in **Analysis II**, e.g. Banach's fixed point theorem. For instance, one can endow $C^{\circ}_{pw}([a,b])$ with the (semi-)norm

$$\|f\|_{L^1} := \int_a^b |f| \,.$$

Then, by considering the aforementioned sequence $(f_n) \subseteq C^{\circ}_{pw}([a,b])$, one can easily show that (f_n) is a Cauchy sequence with respect to $\|\cdot\|_{L^1}$. However, $f_n \to \mathbf{1}_{\mathbb{Q}}$ pointwise. This motivates us to introduce the Banach space $L^1([a,b])$ of integrable functions, and more generally, L^p-spaces later in the course.

Lastly, as we have seen within last term's probability module, measure theory lays below as the foundations for probability theory. As a quick reminder, we recall that a probability space is a special type of measure space and random variables defined on these probability spaces are simply measurable functions to $\mathbb{R}$ (or more exotic fields). This can be interpreted with connotations to real world situations in several ways.

2 Measures and Measure Spaces

2.1 Measures and Measure Spaces

As we would like an adequate theory to assign a notion of "size" on sets, we need to construct a function from a set of sets to $\overline{\mathbb{R}_0^+}$. To achieve this, the natural idea is to construct a function with domain being the power set, however, this is not necessarily always possible or meaningful. Thus, instead of assigning every subset of some set a size, we only look at some collection of "nice" sets.

Definition 2.1 (Algebra)**.** Let X be some set and suppose we denote $\mathcal{X}$ for the power set of X, then a family $\mathcal{A} \subseteq \mathcal{P}(X)$ is called an algebra over X if

- $X \in \mathcal{A}$;
- for all $A \in \mathcal{A}$, $A^c = X \setminus A \in \mathcal{A}$;
- if $(A_k)_{i=1}^n$ is a finite sequence of sets in $\mathcal{A}$, then $\bigcup_{k=1}^n A_k \in \mathcal{A}$.

Definition 2.2 (σ-algebra)**.** Let X be a set, then a σ-algebra $\mathcal{A}$ on X is an algebra on X such that $\mathcal{A}$ is closed under countable unions, i.e. if $(A_k)_{k=1}^\infty$ is a sequence of sets in $\mathcal{A}$, then $\bigcup_{k=1}^\infty A_k \in \mathcal{A}$.

As σ-algebras (and algebras) are simply sets of sets, there is an induced order on σ-algebras by $\subseteq$. If there are two σ-algebras $\mathcal{A}, \mathcal{B}$ such that $\mathcal{A} \subseteq \mathcal{B}$, then we say $\mathcal{A}$ is coarser than $\mathcal{B}$.

Trivially, we find $\{\varnothing, X\}$ is a σ-algebra. Indeed, this is the coarsest σ-algebra. Furthermore, given a set X and a subset $A \subseteq X$, we have $\{\varnothing, A, A^c, X\}$ is also a σ-algebra.

While every σ-algebra is also an algebra, the converse is not true. An counter-example of this is by consider the algebra

$$\mathcal{A} := \{\varnothing\} \cup \left\{ U \mid \exists \bigcup_{k=1}^m (a_k, b_k], m \geq 1, 0 \leq a_k < b_k \leq 1 \right\},$$

on $X = (0,1]$. We see that $\mathcal{A}$ is an algebra (since $(a,b]^c = (0,a] \cup (b,0]$) however $\mathcal{A}$ is not a σ-algebra on X since we can define the sequence $A_k = (0, 1 - 1/k] \in \mathcal{A}$ but $\bigcup_k A_k = (0,1) \notin \mathcal{A}$.

Proposition 1. Let $\mathcal{F}$ be an arbitrary collection of σ-algebra (or algebras) over X. Then the intersections

$$\bigcap \mathcal{F} := \bigcap_{\mathcal{A} \in \mathcal{F}} \mathcal{A},$$

is a σ-algebra (or algebra).

Proof. Straight forward by definition. □

With this, we have a notion of infimum on σ-algebras and hence, we can also define a notion closure.

Definition 2.3 (σ-algebra Generated by a set)**.** Let $\mathcal{C} \subseteq \mathcal{P}(X)$, then

$$\sigma(\mathcal{C}) := \bigcap \{\mathcal{A} \mid \mathcal{C} \subseteq \mathcal{A} \wedge \mathcal{A} \in \mathcal{F}\},$$

where $\mathcal{F}$ is the set of all σ-algebras on X.

As previously shown, $\sigma(\mathcal{C})$ is a intersection of σ-algebras, and so the name suggests, the σ-algebra generated by $\mathcal{C}$ is the smallest σ-algebra containing $\mathcal{C}$. Indeed, we find $\sigma(\varnothing) = \{\varnothing, X\}$ and $\sigma(A) = \{\varnothing, A, A^c, X\}$. Moreover, we see that $\mathcal{C}$ is a σ-algebra if and only if $\sigma(\mathcal{C}) = \mathcal{C}$.

Definition 2.4 (Borel σ-algebra)**.** If $(X, \mathcal{T})$ is a topological space, then the Borel σ-algebra over X is

$$B(X) := \sigma(\mathcal{T}).$$

Unlike topologies, the unions and intersections in a σ-algebra is treated symmetrically.

Proposition 2. If $\mathcal{A}$ is a σ-algebra, then if $(A_k)_{k=1}^{\infty}$ is a sequence of sets in $\mathcal{A}$, then

$$\bigcap_{k=1}^{\infty} A_k \in \mathcal{A}.$$

Proof. By considering de Morgen's identity, we have $\bigcap_{k=1}^{\infty} A_k = (\bigcup_{k=1}^{\infty} A_k^c)^c$. So, since $\bigcup_{k=1}^{\infty} A_k^c \in \mathcal{A}$ as each component is, $\bigcup_{k=1}^{\infty} A_k^c)^c$ and hence $\bigcap_{k=1}^{\infty} A_k$ is also in $\mathcal{A}$. $\square$

Definition 2.5 (Measurable Space)**.** A set X equipped with a σ-algebra $\mathcal{A}$ is called a measurable space and is written as a tuple $(X, \mathcal{A})$. Furthermore, if $A \subseteq X$ is in $\mathcal{A}$, then we say A is a measurable set.

Definition 2.6 (Measure)**.** Let $(X, \mathcal{A})$ is a measurable space. Then a measure on $(X, \mathcal{A})$ is a function $\mu : \mathcal{A} \to [0, \infty]$ such that

- $\mu(\varnothing) = 0$;
- if $(A_k)_{k=1}^{\infty} \subseteq \mathcal{A}$ is a sequence of pairwise disjoint sets, then

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

.

We call the second property σ-additivity.

Definition 2.7 (Measure Space)**.** A measurable space $(X, \mathcal{A})$ equipped with the measure μ is called a measure space and is written as a triplet $(X, \mathcal{A}, \mu)$.

A commonly used measure on any arbitrary measurable space $(X, \mathcal{A})$ is the counting measure μ. As the name suggests, for all $A \in \mathcal{A}$, $\mu(A) = |A|$ if A is finite and ∞ otherwise. Another example of a measure is the Dirac measure $\delta_x : \mathcal{A} \to [0, \infty]$ for some $x \in X$ where for all $A \in \mathcal{A}$, $\delta_x(A) = 1$ if $x \in A$ and 0 otherwise. For the last example, let X be uncountable and let $\mathcal{A} := \{A \subseteq X \mid A \text{ or } A^c \text{ is uncountable}\}$. Then, one can show that $\{X, \mathcal{A}\}$ forms a measurable space and we find that the function $\mu : \mathcal{A} \to [0, \infty]$ defined as $\mu(A) = 0$ if A is countable and $\mu(A) = 1$ if A^c is countable is a measure on $(X, \mathcal{A})$.

Proposition 3. Let $(X, \mathcal{A}, \mu)$ be a measure space. Then,

- if $A, B \in \mathcal{A}$ and $A \subseteq B$, then $\mu(A) \le \mu(B)$;
- if $n \ge 1$, $(A_k)_{k=1}^n$ is a sequence of pairwise disjoint sets in $\mathcal{A}$, then

$$\mu\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n \mu(A_k);$$

- if $(A_k)_{k=1}^\infty$ is a sequence of monotonically increasing sets in $\mathcal{A}$, then

$$\mu\left(\bigcup_{k=1}^\infty A_k\right) = \lim_{k\to\infty} \mu(A_k).$$

We note that the limit in part 3 exists since the limit in monotonically increasing on the extended reals (so bounded by ∞).

- if $(A_k)_{k=1}^\infty$ is a sequence of monotonically decreasing sets in $\mathcal{A}$, if $\mu(A_1) < \infty$, then

$$\mu\left(\bigcap_{k=1}^\infty A_k\right) = \lim_{k\to\infty} \mu(A_k).$$

- if $A \in \mathcal{A}$ and $(A_k)_{k=1}^\infty$ is a sequence in $\mathcal{A}$, then $\mu(A) \le \sum_{k=1}^\infty \mu(A_k)$.

The last property is refereed to as σ-sub-additivity.

Proof. Part 2 is trivial.

(Part 1) As $B = A \sqcup B \setminus A = A \sqcup (A^c \cap B)$ where $A^c \cap B$ is measurable since both A^c and B are. So, by σ-additivity,

$$\mu(B) = \mu(A \sqcup (A^c \cap B)) = \mu(A) + \mu(A^c \cap B) \ge \mu(A).$$

(Part 3) Define $B_1 = A_1$ and $B_{k+1} = A_{k+1} \setminus A_k$. Then, $(B_k)_{k=1}^\infty$ is a sequence of disjoint subset in $\mathcal{A}$. So, by σ-additivity,

$$\mu\left(\bigcup_{k=1}^\infty A_k\right) = \mu\left(\bigcup_{k=1}^\infty B_k\right) = \sum_{k=1}^\infty \mu(B_k) = \lim_{n\to\infty} \sum_{k=1}^n \mu(B_k) = \lim_{n\to\infty} \mu\left(\bigcup_{k=1}^n B_k\right) = \lim_{n\to\infty} \mu(A_k).$$

(Part 4) Define $B_k = A_1 \setminus A_k$, then $(B_k)_{k=1}^\infty$ is a sequence of monotonically increasing sets in $\mathcal{A}$. So, by part 3,

$$\mu\left(\bigcup_{k=1}^\infty B_k\right) = \lim_{k\to\infty} \mu(B_l).$$

Furthermore, as $A \subseteq B$ implies $\mu(B) - \mu(A) = \mu(B \setminus A)$, we have

$$\mu(A_1) - \mu\left(\bigcap_{k=1}^\infty A_k\right) = \mu\left(A_1 \setminus \bigcap_{k=1}^\infty A_k\right) = \mu\left(\bigcup_{k=1}^\infty B_k\right) = \lim_{k\to\infty} \mu(B_k) = \mu(A_1) - \lim_{k\to\infty} \mu(A_k),$$

hence the result.

(Part 5) Exercise. $\square$

2.2 Hahn-Carathéodory Extension Theorem

While it is fun to work with abstract measures in general, it is also useful to consider specific measures. In this section we will look at how to construct measures on arbitrary spaces. The general idea for constructing measures is that, given a *premeasure* $\tilde{\mu}$, we may extend it to an *outer measure* μ^*, and finally with the outer measure, we may restrict it to a measure μ.

Definition 2.8 (Premeasure)**.** Let $\mathcal{A}$ be an algebra on X. A function $\mu : \mathcal{A} \to [0, \infty]$ is a premeasure if it satisfies $\mu(\varnothing) = 0$ and countable additivity; that is, given $(A_n)_{n=1}^{\infty} \subseteq \mathcal{A}$ that is pairwise disjoint,

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

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

Definition 2.9 (Outer Measure)**.** A function $\mu : \mathcal{P}(X) \to [0, \infty]$ is an outer measure if it satisfies $\mu(\varnothing) = 0$ and sub-additivity, that is, if $(A_k)_{k=1}^{\infty}$ is a sequence of subsets of X, and $A \subseteq \bigcup A_k$ then

$$\mu(A) \leq \sum_{k=1}^{\infty} \mu(A_k).$$

Lemma 2.1 (Monoticity of Outer Measures)**.** Let $\mu : \mathcal{P}(X) \to [0, \infty]$ be an outer measure on X. Then for all $A \subseteq B \subseteq X$,

$$\mu(A) \leq \mu(B).$$

Proof. By defining the sequence $A_1 = B$ and $A_n = \varnothing$ for all $n > 1$, by the sub-additivity property of outer measure, we have $A \subseteq B = \bigcup A_n$, and so

$$\mu(A) \leq \sum \mu(A_n) = \mu(B) + \mu(\varnothing) + \cdots = \mu(B).$$

$\square$

In order to prove the aforementioned argument, we will require some proposition. However, as will be demonstrated below, the notion of premeasure is in fact, too strict. Thus, for that reason, let us introduce the following definition.

Definition 2.10 (Cover)**.** A family of sets $\mathcal{K} \subseteq \mathcal{P}(X)$ is called a cover of X if $\varnothing \in \mathcal{K}$ and there exists a sequence $(K_n)_{n=1}^{\infty}$ in $\mathcal{K}$ such that

$$X \subseteq \bigcup_{n=1}^{\infty} K_n.$$

Remark. We note that this is a different definition of covers than that we had learnt for the definition compactness of topological spaces.

Straight away, we see that every algebra $\mathcal{A}$ of X is a cover of X as $X \in \mathcal{A}$ and so we can simply let the the countable sequence be $K_n = X$. Thus, every proposition we prove about covers apply also to algebras.

Proposition 4. Let $\mathcal{K}$ be a cover of X, and $\tilde{\mu} : \mathcal{K} \to [0, \infty]$ be a map such that $\tilde{\mu}(\varnothing) = 0$. Then

$$\mu^* : \mathcal{P}(X) \to [0, \infty] := A \mapsto \inf \left\{ \sum_{j=1}^{\infty} \tilde{\mu}(K_j) \mid (K_j)_{j=1}^{\infty} \subseteq \mathcal{K} \wedge A \subseteq \bigcup_{j=1}^{\infty} K_j \right\}$$

is an outer measure on X.

Proof. μ^* is well defined since for all $A \subseteq X$, by the definition of covers, there exists some $(K_n)_{n=1}^{\infty}$, $\bigcup K_n \supseteq X \supseteq A$, and so, we are taking the infimum of a non-empty set. Clearly $\mu^*(\varnothing) = 0$ since we can choose $K_n = \varnothing$ and so, it remains to show sub-additivity.

Let $(A_n)_{n=1}^{\infty}$ be a sequence of subsets of X, and suppose $A \subseteq \bigcup A_n$. For each $n \in \mathbb{N}$ and $\epsilon > 0$, by the definition of infimum, there exists some $(K_{n,j})_{j=1}^{\infty} \subseteq \mathcal{K}$ such that $A_n \subseteq \bigcup K_{n,j}$ and

$$\sum_{j=1}^{\infty} \tilde{\mu}(K_{n,j}) < \mu^*(A_n) + 2^{-n}\epsilon.$$

Now, by the fact that $A \subseteq \bigcup A_n$, we have $A \subseteq \bigcup_n \bigcup_j K_{n,j}$, and so, again by the definition of μ^*, we have

$$\mu^*(A) \leq \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} \tilde{\mu}(K_{n,j}) < \sum_{n=1}^{\infty} (\mu^*(A_n) + 2^{-n}\epsilon) = \sum_{n=1}^{\infty} \mu^*(A_n) + \epsilon.$$

As $\epsilon > 0$ was arbitrary,

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

so μ^* is a outer measure. □

With that, we can naturally construct a outer measure from a premeasure (and in fact, weaker notions suffices). This, outer measure however, in some sense overshoots measures as it is defined on all subsets of X, and does not attain σ-additivity but only σ-sub-additivity.

Lemma 2.2. Let μ^* be an outer measure on X, then

$$\Sigma := \{A \subseteq X \mid \mu^*(B) = \mu^*(B \cap A) + \mu^*(B \cap A^c)\}$$

is a σ-algebra on X.

Remark. By sub-additivity, we see that Σ is equivalently defined by requiring $\mu^*(B) \geq \mu^*(B \cap A) + \mu^*(B \cap A^c)$.

Proof. $\varnothing \in \Sigma$ since for all $B \subseteq X$, $B \cap \varnothing = \varnothing$ and $B \cap \varnothing^c = B$ and so

$$\mu^*(B \cap \varnothing) + \mu^*(B \cap \varnothing^c) = \mu^*(\varnothing) + \mu^*(B) = \mu^*(B).$$

Let $A \in \Sigma$, then for all $B \subseteq X$,

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

and so, $A^c \in \Sigma$.

To show that Σ is closed under union, let us first show that Σ is closed under finite intersections. Let $A_1, A_2 \in \Sigma$, then for all $B \subseteq X$,

$$\begin{aligned}
\mu^*(B) &= \mu^*(B \cap A_1) + \mu^*(B \cap A_1^c) \\
&= \mu^*(B \cap A_1 \cap A_2) + \mu^*(B \cap A_1 \cap A_2^c) + \mu^*(B \cap A_1^c) \\
&= \mu^*(B \cap (A_1 \cap A_2)) + \mu^*(B \cap (A_1 \cap A_2)^c \cap A_1) + \mu^*(B \cap (A_1 \cap A_2)^c \cap A_1^c) \\
&= \mu^*(B \cap (A_1 \cap A_2)) + \mu^*(B \cap (A_1 \cap A_2)^c),
\end{aligned}$$

and so $A_1 \cap A_2 \in \Sigma$. Now, suppose (A_n) is a sequence of disjoint sets in Σ, it suffices to show $\bigcup_{n=1}^\infty A_n \in \Sigma$ (we can assume this since given a sequence of sets, we may define another disjoint sequence such that their union are equal). Since the sequence is pairwise disjoint, we have

$$\begin{aligned}
\mu^*(B) &= \mu^*(B \cap A_1) + \mu^*(B \cap A_1^c) \\
&= \mu^*(B \cap A_1) + \mu^*(B \cap A_1^c \cap A_2) + \mu^*(B \cap A_1^c \cap A_2^c).
\end{aligned}$$

Indeed, as $A_1 \cap A_2 = \varnothing$, $A_1^c \cap A_2 = A_2$ and so,

$$\begin{aligned}
\mu^*(B) &= \mu^*(B \cap A_1) + \mu^*(B \cap A_2) + \mu^*(B \cap A_1^c \cap A_2^c) \\
&= \cdots \\
&= \sum_{j=1}^n \mu^*(B \cap A_j) + \mu^*\left(B \cap \bigcap_{j=1}^n A_j^c\right) \\
&= \sum_{j=1}^n \mu^*(B \cap A_j) + \mu^*\left(B \cap \overline{\bigcup_{j=1}^n A_j}\right)
\end{aligned}$$

for all $n \in \mathbb{N}$. By sub-additivity, we have

$$\mu^*\left(B \cap \bigcup_{j=1}^n A_j\right) \leq \sum_{j=1}^n \mu^*(B \cap A_j),$$

and so, by taking $n \to \infty$, we have

$$\mu^*(B) \geq \mu^*\left(B \cap \bigcup_{j=1}^\infty A_j\right) + \mu^*\left(B \cap \overline{\bigcup_{j=1}^\infty A_j}\right),$$

implying $\bigcup A_n \in \Sigma$. $\square$

Theorem 1 (Hahn-Carathéodory Extension Theorem)**.** Let X be an arbitrary set, $\mathcal{A}$ an algebra over X and $\tilde{\mu} : \mathcal{A} \to [0, \infty]$ a premeasure on X. Then by defining μ^* as in proposition 4, Σ as in lemma 2.2, and by defining $\mu = \mu^* |_\Sigma$, then,

- (X, Σ, μ) is a measure space;
- $\mathcal{A} \subseteq \Sigma$;

- for all $A \in \mathcal{A}$, $\mu(A) := \mu^*(A) = \tilde{\mu}(A)$.

Proof. By construction, we have μ^* is an outer measure on X and Σ is a σ-algebra on X, and so, for the first part of the proof, it remains to show that $\mu = \mu^* \mid_\Sigma$ is a measure on the measurable space (X, Σ).

Indeed, $\mu(\varnothing) = 0$ as $\mu^*(\varnothing) = 0$ since it is an outer measure, so let us consider σ-additivity. Let $A, B \in \Sigma$, $A \cap B = \varnothing$, then

$$\begin{aligned}\mu(A \cup B) = \mu^*(A \cup B) &= \mu^*((A \cup B) \cap A) + \mu^*((A \cup B) \cap A^C) \\ &= \mu^*(A) + \mu^*(B),\end{aligned}$$

and so, by induction μ is finitely additive. So, if $(A_n)_{n=1}^\infty$ is a sequence of pairwise disjoint measurable sets, then, for all $m \in \mathbb{N}$,

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

where the inequality is due to the monotonicity of outer measures. So, by taking $m \to \infty$ we have

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

However, by the sub-additivity of outer measures,

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

and so $\sum \mu(A_n) \geq \mu\left(\bigcup A_n\right)$ resulting in μ being a measure.

For the second part of the theorem, it suffices to show that for all $A \in \mathcal{A}$, for all $B \subseteq X$,

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

By recalling the definition of μ^*, for all $\epsilon > 0$, there exists some sequence $(K_j)_{j=1}^\infty \subseteq \mathcal{A}$, $B \subseteq \bigcup K_j$,

$$\sum_{j=1}^\infty \tilde{\mu}(K_j) \leq \mu^*(B) + \epsilon.$$

Now, as $B \subseteq \bigcup K_j$, $B \cap A \subseteq \bigcup (K_j \cap A)$ and $B \cap A^c \subseteq \bigcup (K_j \cap A^c)$, and so, by sub-additivity

$$\begin{aligned}\mu^*(B \cap A) + \mu^*(B \cap A^c) &\leq \sum \mu^*(K_j \cap A) + \sum \mu^*(K_j \cap A^c) \\ &= \sum (\tilde{\mu}(K_j \cap A) + \tilde{\mu}(K_j \cap A^c)) \\ &= \sum \tilde{\mu}(K_j) \leq \mu^*(B) + \epsilon,\end{aligned}$$

where we used the last part of the theorem to exchange μ^* for $\tilde{\mu}$. Thus, as $\epsilon > 0$ was arbitrary, the inequality follows.

Lastly, to show that μ^* agrees with $\tilde{\mu}$ on $\mathcal{A}$, let $A \in \mathcal{A}$. By the monotonicity of outer measures, we have $\mu^*(A) \geq \tilde{\mu}(A)$ and so, it suffices to show there exists some $(K_j) \subseteq \mathcal{A}$ such that $A \subseteq \bigcup K_j$ and $\sum \tilde{\mu}(K_j) = \tilde{\mu}(A)$. But, this is trivial by simply choosing $A_1 = A$ and $A_n = \varnothing$ for all $n > 1$ and so, we are done! $\square$

With the Hahn-Carathéodory extension theorem, we can obtain a measure (and also a measurable space), just from a premeasure on some algebra. This is a very useful theorem as algebras and premeasures are rather weak notions and can be easily constructed through many means. However, this theorem does not yet guarantee the uniqueness, so one might ask whether or not there are other extensions. We shall take a look at that now.

Definition 2.11 (σ-Finite)**.** A premeasure $\tilde{\mu}$ on the algebra $\mathcal{A}$ on X is σ-finite if there exists disjoint sets $(S_n)_{n=1}^{\infty} \subseteq \mathcal{A}$ such that $X = \bigcup_{n=1}^{\infty} S_n$ and $\tilde{\mu}(S_n) < \infty$ for all $n \geq 1$.

Theorem 2 (Uniqueness of the Hahn-Carathéodory Extension)**.** Under the assumption of the Hahn-Carathéodory extension theorem, and if $\tilde{\mu}$ is σ-finite. Then, for all outer measures $\nu : \mathcal{P}(X) \to [0, \infty]$,

$$\nu \mid_{\Sigma} = \mu,$$

where μ is the outer measure obtained from the Hahn-Carathéodory extension.

Proof. Let $A \in \Sigma$, and we will first show $\nu(A) \leq \mu(A)$. Let $(A_n)_{n=1}^{\infty} \subseteq \mathcal{A}$ be some sequence such that $A \subseteq \bigcup A_n$. Then, by the sub-additivity of ν,

$$\nu(A) \leq \sum \nu(A_n) = \sum \tilde{\mu}(A_n).$$

So, by taking the infimum over all possible (A_n), the inequality is achieved.

To show the reverse inequality, let us suppose $(*)$ that there exists some $S \in \mathcal{A}$ such that $A \subseteq S$ and $\mu(S) < \infty$. Then,

$$\nu(A) + \nu(S \cap A^c) \leq \mu(A) + \mu(S \cap A^c) = \mu(S) = \tilde{\mu}(S) = \nu(S).$$

Now, by sub-additivity,

$$\nu(S) \leq \nu(A) + \nu(S \cap A^c),$$

and so, $\nu(A) + \nu(S \cap A^c) \geq \mu(A) + \mu(S \cap A^c)$, hence[1],

$$\nu(A) - \mu(A) \geq \mu(S \cap A^c) - \nu(S \cap A^c) \geq 0,$$

as $\mu \geq \nu$ as shown previously. Thus, $\nu(A) \geq \mu(A)$ and so, $\nu \mid_{\Sigma} = \mu$.

Now, to relax the assumption that such S exists, we shall use the σ-finiteness of $\tilde{\mu}$. Let $(S_n) \subseteq \mathcal{A}$ be the sequence as described by the σ-finiteness of $\tilde{\mu}$, and define $A_n = A \cap S_n$. Then, by construction, the sets A_n are pairwise disjoint and $A = \bigcup A_n$. Since $\mu = \nu$ on sets, which $(*)$ is satisfied, we have

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

[1] Wlog. $\mu(A) < \infty$ since if otherwise, $\infty < \nu(A) + \nu(S \cap A^c)$. Since $\nu(S \cap A^c) \leq \nu(S) \leq \mu(S) < \infty$, $\nu(A) = \infty$ and hence $\nu(A) \geq \mu(A)$.

for all $m \geq 1$ (since we can choose $S = \bigcup_{n=1}^{m} S_n$). Now, by monotonicity,

$$\nu(A) = \nu\left(\bigcup_{n=1}^{\infty} A_n\right) \geq \nu\left(\bigcup_{n=1}^{m} A_n\right) = \mu\left(\bigcup_{n=1}^{m} A_n\right),$$

for all m, the inequality is achieved by taking $m \to \infty$. $\square$

2.3 The Lebesgue Measure

With the Hahn-Carathéodory extension theorem in mind, we may construct the Lebesgue measure on Euclidean spaces. As with the proof of the extension theorem, we shall construct a premeasure assigning $b - a$ to every interval (a, b) and then, extend that to a general measure.

Definition 2.12. For $a = (a_1, \cdots, a_n) \in \mathbb{R}^n$, $b = (b_1, \cdots, b_n) \in \mathbb{R}^n$, we define the interval between a and b to be the set

$$(a, b) := \prod_{k=1}^{n} (a_k, b_k) = \{(x_1, \cdots, x_n) \mid a_k < x_k < b_k \;\; 1 \leq k \leq n\},$$

if $a_k < b_k$ for all k and $(a, b) = \varnothing$ otherwise.

Similarly, we define the half open and closed intervals in $\mathbb{R}^n$ with $(a, b] := \prod_{k=1}^{n}(a_k, b_k]$, $[a, b) := \prod_{k=1}^{n}[a_k, b_k)$ and $[a, b] := \prod_{k=1}^{n}[a_k, b_k]$. Indeed, we also allow $\pm\infty$ as an endpoint and we shall from this point forward refer to such sets as intervals in $\mathbb{R}^n$.

Definition 2.13 (Elementary Figure)**.** A set $I \subseteq \mathbb{R}^n$ is an elementary figure if its the union of finitely many disjoint intervals; and in this section, we shall denote $\mathcal{A} \subseteq \mathcal{P}(\mathbb{R}^n)$ for the set of elementary figures in $\mathbb{R}^n$.

We see that $\mathcal{A}$ is an algebra on $X = \mathbb{R}^n$ and so, we may define a premeasure on $\mathcal{A}$. Let $\tilde{\lambda} : \mathcal{A} \to [0, \infty]$ be the function such that for all $a, b \in \mathbb{R}^n$,

$$\tilde{\lambda}((a, b)) = \tilde{\lambda}((a, b]) = \tilde{\lambda}([a, b)) = \tilde{\lambda}([a, b]) := \prod_{k=1}^{n} (b_k - a_k),$$

if $a_k < b_k$ and 0 otherwise. Furthermore, for $\bigcup_{k=1}^{m} I_k \in \mathcal{A}$,

$$\tilde{\lambda}\left(\bigcup_{k=1}^{m} I_k\right) = \sum_{k=1}^{m} \tilde{\lambda}(I_k),$$

where $I_1, \cdots, I_k$ are disjoint intervals.

Lemma 2.3. The map $\tilde{\lambda}$ is a premeasure on $\mathcal{A}$.

Proof. As it is trivially true that $\tilde{\lambda}(\varnothing) = 0$ to show that $\tilde{\lambda}$ is a premeasure, it suffices to show σ-additivity within $\mathcal{A}$.

We note that $\tilde{\lambda}$ is finitely additive by definition, hence monotone on $\mathcal{A}$, and therefore if $I = \bigcup_{k=1}^{\infty}$ where I_k are disjoint intervals,

$$\tilde{\lambda}(I) \geq \tilde{\lambda}\left(\bigcup_{k=1}^{m} I_k\right) = \sum_{k=1}^{m} \tilde{\lambda}(I_k),$$

for all $m \in \mathbb{N}$. Thus, by taking $m \to \infty$,

$$\tilde{\lambda}(I) \geq \sum_{k=1}^{\infty} \tilde{\lambda}(I_k).$$

For the reverse inequality, we use the what is called a *compactness argument* to reduce to finite additivity. Wlog. we may assume $\sum \tilde{\lambda}(I_k) < \infty$ (since if otherwise the reverse inequality is trivial) and that I is a single interval with end points a, b (since if the inequality is true for a single interval, it is also true for the sum of finitely many intervals).

Let $\bar{I}$ be the closure of I and for all $L > 0$ we define $\bar{I}_L := \bar{I} \cap [-L, L]^n$. By Heine-Borel, $\bar{I}_L$ is compact and moreover, by taking $L \to \infty$,

$$\tilde{\lambda}(\bar{I}_L) \to \tilde{\lambda}(\bar{I}) = \tilde{\lambda}(I).$$

Now, for all intervals J with end points α, β, we define the interval $J^\epsilon \supseteq J$ with endpoints $\alpha^\epsilon \neq \alpha$, $\beta^\epsilon \neq \beta$ such that

$$\tilde{\lambda}(J^\epsilon) \leq (1+\epsilon)^n \tilde{\lambda}(J).$$

Lastly, for all $k \in \mathbb{N}$, we define the open intervals $\tilde{I}_k$ with $\tilde{I}_k \supseteq I_k^\epsilon$ satisfying

$$\tilde{\lambda}(\tilde{I}_k) < (1+\epsilon)^n \lambda(I_k) + \epsilon 2^{-k}.$$

So $\bar{I}_L \subseteq \bar{I} \subseteq \bigcup_{k=1}^{\infty} I_k^\epsilon \subseteq \bigcup_{k=1}^{\infty} \tilde{I}_k$, and $\{\tilde{I}_k\}$ forms an open cover of $\bar{I}_L$, and hence, by compactness, there exists a finite subcover for $\bar{I}_L$, that is, there exists some m such that (by reordering), $\bar{I}_L \subseteq \bigcup_{k=1}^{m} I_k$. It follows that

$$\tilde{\lambda}(\bar{I}_L) \leq \tilde{\lambda}\left(\bigcup_{k=1}^{m} \tilde{I}_k\right) \leq \sum_{k=1}^{m} \tilde{\lambda}(\tilde{I}_k) \leq (1+\epsilon)^n \sum_{k=1}^{\infty} \lambda(I_k) + \epsilon.$$

Thus, by taking $\epsilon \to 0$ and then $L \to \infty$ we have the reverse inequality and so $\tilde{\lambda}$ is a premeasure on $\mathcal{A}$. $\square$

With this lemma, one can immediately apply the Hahn-Carathéodory extension theorem resulting in the Lebesgue measure on Euclidean spaces, and furthermore, by considering $\tilde{\lambda}((z, z+1]) = 1 < \infty$, and $\bigcup_{z \in \mathbb{Z}^n} (z, z+1] = \mathbb{R}^n$, we have $\tilde{\lambda}$ is σ-finite, and hence, the Lebesgue measure is unique.

Lemma 2.4. Let $\mathcal{B}(\mathbb{R}^n)$ be the Borel σ-algebra on $\mathbb{R}^n$, then $\mathcal{B}(\mathbb{R}^n) \subseteq \Sigma$ where Σ is the σ-algebra induced by the Lebesgue measure.

Proof. Since $\mathcal{B}(\mathbb{R}^n)$ is the σ-algebra generated by the set of open sets of $\mathbb{R}^n$, it suffices to show that for all open sets $O \subseteq \mathbb{R}^n$, $O \in \sigma(\mathcal{A}) \subseteq \Sigma$.

For all $m \in \mathbb{N}$, define
$$C_m := \{[z, z + 2^{-m}) \mid z \in 2^{-m}\mathbb{Z}^n\} \subseteq \mathcal{A},$$
that is the grid of half open cubes covering $\mathbb{R}^n$ with individual cubes having length 2^{-m}. Then, by letting $C'_m := \{U \in C_m \mid U \subseteq O\}$, we have $C = \bigcup_{m\in\mathbb{N}} C'_m$ which is a countable set of half open cubes. Now, we see that $C = O$ since $C \subseteq O$ trivially and for all $o \in O$, as O is open, there exists some $\epsilon > 0$ such that $B_\epsilon(o) \subseteq O$ and so, o is contained in one of the cubes with length $< 2^{-m}$ where $m > 2/\epsilon$ and so $O \in \sigma(\mathcal{A}) \subseteq \Sigma$. $\square$

With that, we see that all Borel sets in $\mathbb{R}^n$ are Lebesgue measurable and restricting $\tilde{\lambda}$ onto $\mathcal{B}(\mathbb{R}^n)$, we have the following measure space on $\mathbb{R}^n$.

Definition 2.14. The Lebesgue measure $\lambda : \mathcal{B}(\mathbb{R}^n) \to [0, \infty]$ is the Hahn-Carathéodory extension of $\tilde{\lambda}$ restricted onto $\mathcal{B}(\mathbb{R}^n)$.

We note that, if $(X, \mathcal{F}, \mu)$ is a measure space and $A \in \mathcal{F}$, then by defining $\mathcal{F} \mid_A := \{A \cap B \mid B \in \mathcal{F}\}$ and $\mu \mid_A (B) := \mu(B)$ for all $B \in \mathcal{F}$, $B \subseteq A$, it is easy to see that $\mathcal{F} \mid_A$ is a σ-algebra on A and $\mu \mid_A$ is a measure on $(A, \mathcal{F} \mid_A)$ and is called the restriction of μ to A. Indeed, with this in mind, we see that the Lebesgue measure can be restricted on small sets such as intervals; in particular, by restricting the Lebesgue measure on $[0, 1]$, the resulting measure $\lambda \mid_{[0,1]}$ is a probability measure.

2.4 Lebesgue Measurable Sets

We investigate which real sets are Lebesgue measurable.

Proposition 5. For all $A \in \mathcal{B}(\mathbb{R}^n)$,
$$\lambda(A) = \inf_{G \supseteq A; G\text{open}} \lambda(G).$$

Proof. Since measures are monotone, $(\leq)$ is established. By recalling the construction of the Lebesgue measure, and by the properties of inf, for all $\epsilon > 0$, there exists some $(K_n)_{n=1}^\infty \subseteq \mathcal{A}$, such that $A \subseteq \bigcup K_n$ and
$$\lambda(A) + \epsilon = \lambda^*(A) + \epsilon \geq \sum_{n=1}^\infty \tilde{\lambda}(K_n) \geq \tilde{\lambda}\left(\bigcup K_n\right) = \lambda\left(\bigcup K_n\right) \geq \inf_{G \supseteq A; G\text{open}} \lambda(G).$$
since the union of open sets is open, $\bigcup K_n$ is open and contains A. So, as $\epsilon > 0$ was arbitrary, $(\geq)$ is established. $\square$

By inspection of the proof, we find the statement to be true for any real sets provided with change λ to the outer measure λ^* on the left hand side. This *regularity* of measurable sets is expressed in the following.

Proposition 6. If $A \in \mathcal{B}(\mathbb{R}^n)$, then for all $\epsilon > 0$, there exists some $G \supseteq A$, G open such that $\lambda(G \setminus A) < \epsilon$.

Proof. One can of course prove it using the construction of the Lebesgue measure, however, a stronger proposition holds. $\square$

Proposition 7. Let $\mathcal{A}$ be an algebra and μ an measure on $\sigma(\mathcal{A})$ which is σ-finite on $\mathcal{A}$. Then, for all $A \in \sigma(\mathcal{A})$, and $\epsilon > 0$, there exist disjoint sets $(A_n)_{n=1}^{\infty} \in \mathcal{A}$ such that $A \subseteq \bigcup A_n$ and $\mu\left(\bigcup A_n \setminus A\right) < \epsilon$.

Proof. See problem sheet 4. $\square$

Furthermore, by applying proposition 6 to A^c for some $A \in \mathcal{B}(\mathbb{R}^n)$, there exists some $\tilde{G}$ open, containing A^c such that $\lambda(\tilde{G} \setminus A^c) < \epsilon$ for all $\epsilon > 0$. Then, by defining $F := \tilde{G}^c$, we have $F \subseteq A$ closed, such that

$$\lambda(A \setminus F) = \lambda(A \cap F^c) = \lambda(A \cap \tilde{G}) = \lambda(\tilde{G} \setminus A^c) < \epsilon.$$

Thus,

$$\lambda(G \setminus F) = (\lambda(G) - \lambda(A)) + (\lambda(A) - \lambda(F)) = \lambda(G \setminus A) + \lambda(A \setminus F) < 2\epsilon.$$

So, for all $A \in \mathcal{B}(\mathbb{R}^n)$, there exists some open $G \supseteq A$ and some closed $F \subseteq A$ such that $\lambda(G \setminus F) < \epsilon$ for all $\epsilon > 0$.

Proposition 8 (Transitional Invariance of λ)**.** Let $\Phi_{x_0} : \mathbb{R}^n \to \mathbb{R}^n : x \mapsto x + x_0$ for some $x_0 \in \mathbb{R}^n$. Then,

$$\lambda(\Phi_{x_0}(A)) = \lambda(A),$$

for all $A \in \mathcal{B}(\mathbb{R}^n)$.

Proof. Since, for all $A \in \mathcal{B}(\mathbb{R}^n)$, there exists some open $G \supseteq A$ such that $\lambda(G) - \lambda(A) < \epsilon$, it suffices to show transitional invariance for open sets. Now, as shown previously, any open sets in $\mathbb{R}^n$ can be written as a disjoint union of countable intervals, the result follows since Lebesgue measures are transitional invariant on intervals by definition. $\square$

Proposition 9. $\mathcal{B}(\mathbb{R}) \neq \mathcal{P}(\mathbb{R})$.

Proof. As we have seen last term, the Vitali set is a classical example of a non-measurable real set.

Define the equivalence relation $\sim$ such that, for all $x, y \in (0, 1]$, $x \sim y \iff x - y \in \mathbb{Q}$. Then, with axiom of choice, for each equivalence class $[x] \in (0, 1] \setminus \sim$, we choose exactly one $v \in [x]$ and we define the Vitali set V to be the set of these choices.

Now, let $A := \mathbb{Q} \cap (-1, 1]$ and then,

$$(0, 1] \subseteq \bigcup_{q \in A} (q + V) \subseteq [-1, 2],$$

where the first inclusion is true since, for all $x \in (0, 1]$, x belongs to an equivalence class $[x']$ where $x' \in V$. Now, as $x \in [x']$ implies $x - x' \in \mathbb{Q}$, we can simply choose $q = x - x'$ and so $x = q + x' \in \bigcup q + V$.

Thus, if V is measurable, then $1 \leq \lambda(\bigcup(q+V)) \leq 3$. Now, by observing that for all $p, q \in \mathbb{Q}$, if $p \neq q$ then $p + V \cap q + V = \varnothing$,

$$\lambda\left(\bigcup_{q \in A}(q+V)\right) = \sum_{q \in A} \lambda(q+V) = \sum_{q \in A} \lambda(V),$$

since λ is transitional invariant. However, as $\lambda(V)$ is bounded above by 3, $\lambda(V) = 0 \#$ as $\lambda(V) \geq 1$.

So $V \notin \mathcal{B}(\mathbb{R})$ and hence $\mathcal{B}(\mathbb{R}) \neq \mathcal{P}(\mathbb{R})$. $\square$