I. Construction of Lebesgue Measure

1. Length of Open Subsets of $\mathbb{R}$

Recall 1.1. **Equivalence Class**
Let X be a nonempty set and $\sim$ be a relation on X. We say $\sim$ is an ***equivalence relation*** if for every $x, y, z \in X$,

(a) $x \sim x$; *reflexivity*

(b) $x \sim y$ implies $y \sim x$; and *symmetry*

(c) $x \sim y, y \sim z$ imples $x \sim z$. *transitivity*

Let X be a nonempty set and let $\sim$ be an equivalence relation on X. This gives rise to a decomposition

$$X = \bigcup_{i \in I} C_i$$

where $\{C_i\}_{i \in I}$ is a disjoint collection of subsets of X, and such that for all $x, y \in X$,

(a) if $x, y \in C_i$ for some $i \in I$, then $x \sim y$; and

(b) if $x \in C_i, y \in C_j$ for some distinct $i, j \in I$, then $x \nsim y$.

Recall 1.2. **Equivalence Class** of an Equivalence Relation
Consider the above setting. We call each C_i an ***equivalence class*** of $\sim$.

Conversely, if we have a partition $\{C_i\}_{i \in I}$ of a nonempty set X, then we can define an equivalence relation $\sim$ on X as follows: given $x, y \in X$,

$$x \sim y \iff \exists i \in I \left[x, y \in C_i\right]. \quad [1.1]$$

We next look at how the notion of equivalence relation goes into the material of today's class.

Exercise 1.1.

Let A be a nonempty subset of $\mathbb{R}$. Define $\sim$ on A as follows: for every $x, y \in A$ (say $x \leq y$ without loss of generality),

$$x \sim y \iff [x, y] \subseteq A. \quad [1.2]$$

(a) Prove that $\sim$ is an equivalence relation on A.

(b) Let $A = \bigcup_{i \in I} C_i$ be the decomposition into equivalence classes for the above relation $\sim$. Prove that every C_i is an interval.

Proof.

(a) Let $x, y, z \in A$. It is clear that $\sim$ is reflexive and symmetric. To show that $\sim$ is transitive, suppose $x \sim y, y \sim z$. We break down into few cases.

If $x \leq y \leq z$, then $[x, z] = [x, y] \cup [y, z] \subseteq A$.

If $x \leq z \leq y$, then $[x, z] \subseteq [x, y] \subseteq A$.

Other cases can be verified in a similar manner.

(b) Suppose C_i is not an interval. Then there exists $x, y, z \in A$ such that $x < y < z$ and $x, z \in C_i$ but $y \notin C_i$. But this means $[x, z]$ is not contained in C_i, which is a contradiction.

QED

In the framework of Exercise 1.1, we will refer to the equivalence classes of ~ by calling them *interval components* of A.

A silly question that comes to mind: what is a convenient definition, to be used in the solution to Exercise 1.1(b), for the notion of interval? It is preferable to go with the following unified description: *a set $J \subseteq R$ is said to be an interval when it has the property that, if $x \leq y \leq z$ with $x, z \in J$,then $y \in J$.*

Def'n 1.3. **Open** Subset of $\mathbb{R}$

A subset $A \subseteq \mathbb{R}$ is ***open*** if for every $x \in A$, there exists $r > 0$ such that $(x - r, x + r) \subseteq A$.

Proposition 1.1.

Let $A \subseteq \mathbb{R}$ be nonempty and open and let $A = \bigcup_{i \in I} C_i$ be the decomposition of A into interval components. Then every C_i is an open interval.

Proof. Fix $i \in I$ for which we will prove that C_i is open (we know that C_i is an interval from Exercise 1.1(b)). To that end, let us also fix a point $x \in C_i$. We need to find $r > 0$ such that $(x - r, x + r) \subseteq C_i$.

Since $x \in A$ and A is open, there exists $r > 0$ such that $(x - r, x + r) \subseteq A$. We will prove that this r is what we need – that is, we can strengthen the inclusion $(x - r, x + r) \subseteq A$ to $(x - r, x + r) \subseteq C_i$.

Choose a point $y \in (x - r, x + r)$, for which we have to check $y \in C_i$. We argue like this

$$\begin{aligned} y \in (x - r, x + r) &\implies \text{the whole interval with endpoints at } x, y \text{ is contained in } (x - r, x + r) \\ &\implies \text{the whole interval with endpoints at } x, y \text{ is contained in } A \\ &\implies x \sim y \\ &\implies y \in C_i. \end{aligned}$$

QED

Let us summarize what is the decomposition into interval components for a nonempty open $A \subseteq \mathbb{R}$: it is a decomposition $A = \bigcup_{i \in I} C_i$, where

(a) every C_i is a nonempty open interval;

(b) $C_i \cap C_j = \varnothing$ for all distinct $i, j \in I$; and

(c) if $x, y \in A$ (without loss of generality, $x < y$) belong to different component intervals, then there is $z \in \mathbb{R} \setminus A$ such that $x < z < y$.

It is instructive to look a bit more detail at the condition (c). This was written in a way which simply stated the fact that if $x \in C_i, y \in C_j$ for some distinct $i, j \in I$, then $x \nsim y$. It is easy to check that (c) can be rephrased in a strong form, as follows.

Pick distinct indices $i, j \in I$ and pick two points $x_0 \in C_i, y_0 \in C_j$. Without loss of generality say $x_0 < y_0$. Then there exists $z \in \mathbb{R} \setminus A$ such that $x < z$ for all $x \in C_i$ and $z < y$ for all $y \in C_j$.

What the above condition says is that the point z *separates* C_i from C_j in the stronger sense that all of C_i is to the *left* of z while all of C_j is to the *right* of z.

Here is another useful fact about (c): it actually follows for free if we have (a), (b). This is formally stated in the next exercise.

Exercise 1.2.

Let $A \subseteq \mathbb{R}$ be nonempty and open and suppose we are given a decomposition $A = \bigcup_{i \in I} C_i$, where every C_i is a nonempty open interval and $C_i \cap C_j = \varnothing$ for all distinct $i, j \in I$.

(a) Let $x, y \in A$ with $x < y$ and suppose that $x \in C_i, y \in C_j$ with $i \neq j$. Prove that there exists $z \in \mathbb{R} \setminus A$ such that $x < z < y$.

(b) Prove that the intervals C_i are precisely the equivalence classes for the equivalence relation ~ defined in [1.2].

Proof.

(a) Write $C_i = (a_i, b_i)$, $C_j = (a_j, b_j)$ for some $a_i, a_j, b_i, b_j \in \mathbb{R}$. Now note that

$$\varnothing = C_i \cap C_j = \left\{\max\left\{a_i, a_j\right\}, \min\left\{b_i, b_j\right\}\right\}.$$

This means $\max\left\{a_i, a_j\right\} > \min\left\{b_i, b_j\right\}$. But we know that $a_i < x < y < b_j, a_i < b_i, a_j < b_j$. Hence we conclude that $b_i < a_j$.

Now the interval $\left[b_i, a_j\right]$ is a nonempty closed interval, so cannot be written as a union of open intervals (since union of open sets is open, and the only clopen sets in $\mathbb{R}$ are $\varnothing, \mathbb{R}$; from PMATH 351, we know a metric space X is connected if and only if $\varnothing, X$ are the only clopen sets and $\mathbb{R}$ is a connected space). Hence there exists $z \in \left[b_i, a_j\right] \setminus A$, and by construction

$$x < b_i \leq z \leq a_j < y,$$

as required.

(b) Suppose $x, y \in C_i$ (with $x < y$ without loss of generality) for some $i \in I$. Since C_i is an interval, C_i is convex, so $[x, y] \subseteq C_i \subseteq A$. Hence $x \sim y$.

Conversely, suppose $x, y \in A$ are such that $x \sim y$ but $x \in C_i, y \in C_j$ for some distinct $i, j \in I$ for contradiction. Then by (a), we know there is $z \in \mathbb{R} \setminus A$ such that $x < z < y$. But $x \sim y$ if and only if $[x, y] \subseteq A$ and $z \in [x, y]$. This is a contradiction.

QED

Exercise 1.3.

Consider nonempty open $A \subseteq \mathbb{R}$ and the decomposition $A = \bigcup_{i \in I} C_i$ of A into interval components. Prove that I is countable.

Proof. For each $i \in I$, choose $q_i \in C_i$ such that $q_i \in \mathbb{Q}$. This is possible since $\mathbb{Q}$ is dense in $\mathbb{R}$ and every C_i is a nonempty *open* interval, so that $\mathbb{Q} \cap C_i \neq \varnothing$ for all $i \in I$. Since Q_i's are disjoint, this defines an injection $\phi : I \to \mathbb{Q}$ by

$$\phi(i) = q_i$$

for all $i \in I$. Hence

$$|I| \leq |\mathbb{Q}| = \aleph_0.$$

QED

The notion of length for an open subset of $\mathbb{R}$ is now easy to define, since we have a clear idea of what should be the length of an open *interval*, and we can use the structural result from Proposition 1.1.

Def'n 1.4. **Length** of an Open Subset of $\mathbb{R}$

Let $A \subseteq \mathbb{R}$ be open. We define a quantity $\lambda(A) \in [0, \infty]$, which we will call (for now) ***length*** of A, as follows.

(a) If $A = \varnothing$, then $\lambda(A) = 0$.

(b) Suppose $A \neq \varnothing$ and consider the decomposition $A = \bigcup_{i \in I} C_i$ into interval components.

  (i) If there is $i \in I$ such that C_i is unbounded, then $\lambda(A) = \infty$.

  (ii) If every C_i is bounded, say $C_i = (a_i, b_i)$ for some $a_i, b_i \in \mathbb{R}$. Then we define

$$\lambda(A) = \sum_{i \in I} (b_i - a_i) \in [0, \infty]. \tag{1.3}$$

Here is a little discussion around the meaining of the sum in [1.3]. Note that, due to Exercise 1.3, the index set I can be re-denoted in a way which makes it that either $I = \mathbb{N}$ or $I = \{1, \ldots, n\}$ for some $k \in \mathbb{N}$. In other words, we are dealing either with a finite sum, or with a series of nonnegative real numbers (for which the order of summation does not matter).

There actually is a way to handle the sum [1.3] which does not require a discussion around the cardinality of I, and is simply based on the notion of *supremum* for a subset of $[0, \infty]$. More precisely, given open $A \subseteq \mathbb{R}$ which falls in (c) of Def'n 1.4, it may sometimes be convenient to rewrite the formula [1.3] in the form

$$\lambda(A) = \sup\left\{\sum_{j=1}^{n} b_j - a_j : n \in \mathbb{N}, (a_1, b_1), \ldots, (a_n, b_n) \text{ are pairwise disjoint interval components of } A\right\} \tag{1.4}$$

If A is just a bounded open interval, say $A = (a, b)$, then we come to the unsurprising conclusion:

$$\lambda((a, b)) = b - a.$$

A consequence of this is yet another way of writing [1.3]:

$$\lambda(A) = \sum_{i \in I} \lambda(C_i). \qquad [1.5]$$

It is useful to generalze the latter formula to a situation where the sets on the right-hand side don't have to be intervals.

Proposition 1.2.
Let $\{A_j\}_{j \in J}$ be a collection of pairwise disjoint open subsets of $\mathbb{R}$. Then

$$\lambda\left(\bigcup_{j \in J} A_J\right) = \sum_{j \in J} \lambda(A_j) \in [0, \infty]. \qquad [1.6]$$

Tl;dr

The collection of all open subsets of $\mathbb{R}$,

$$\mathcal{T} = \{A \subseteq \mathbb{R} : A \text{ is open}\}$$

goes under the name of *topology* of $\mathbb{R}$. $\mathcal{T}$ has some good properties in connection to set-operations:

(a) for every $\mathcal{E} \subseteq \mathcal{T}$, $\bigcup \mathcal{E} \in \mathcal{T}$; and *closure under union*

(b) for every finite $\mathcal{E} \subseteq \mathcal{T}$, $\bigcap \mathcal{E} \in \mathcal{T}$. *closure under finite intersection*

Then the association $A \mapsto \lambda(A)$ discussed before can be pitched by saying that: *we have defined a function $\lambda : \mathcal{T} \to [0, \infty]$.*

We can restate the properties of λ.

(a) $\lambda(\varnothing) = 0$.

(b) $\lambda((a, b)) = b - a$ for all $a, b \in \mathbb{R} \cup \{\pm\infty\}$ with $a < b$. *values on open intervals*

(c) If $\{A_n\}_{n=1}^{\infty} \subseteq \mathcal{T}$ is a collection of disjoint sets, then $\lambda\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \lambda(A_n)$. *additivity for disjoint union*

These properties turn out to completely determine λ, as stated by the next exercise.

Exercise 1.4.
Let $\mu : \mathcal{T} \to [0, \infty]$ which has the same properties with λ. That is:

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

(b) $\mu((a, b)) = b - a$ for all $a, b \in \mathbb{R} \cup \{\pm\infty\}$ with $a < b$; and

(c) if $\{A_n\}_{n=1}^{\infty} \subseteq \mathcal{T}$ is a collection of disjoint sets, then $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$.

Prove that $\mu = \lambda$.

Proof. Let $A \in \mathcal{T}$. Since A is open, there exists a countable partition $\{A_i\}_{i \in I}$ of A into open intervals. Then

$$\mu(A) = \mu\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} \mu(A_i) = \sum_{i \in I} \lambda(A_i) = \lambda\left(\bigcup_{i \in I} A_i\right) = \lambda(A).$$

QED

Now that we started on the path of studying $\lambda : \mathcal{T} \to [0, \infty]$, let us look for some other natural properties of λ.

Proposition 1.3.
If $A, B \in \mathcal{T}$ are such that $A \subseteq B$, then $\lambda(A) \leq \lambda(B)$.

Proof? Let $A, B \in \mathcal{T}$ with $A \subseteq A$. We denote $A_1 = A, A_2 = B \setminus A$. Then $A_1 \cup A_2 = B$. It is also clear that $A_1 \cap A_2 = \varnothing$. Hence

$$\lambda(A_1 \cup A_2) = \lambda(A_1) + \lambda(A_2).$$

We can thus write

$$\lambda(B) = \lambda(A_1 \cup A_2) = \lambda(A_1) + \lambda(A_2) \geq \lambda(A_1) = \lambda(A),$$

as required.

???

The issue of the above *proof* is that $\mathcal{T}$ is *not closed under set difference* (so that $\lambda(A_2)$ is *not defined*). That is something we will have to cope with.

2. Positive Measure on a σ-algebra

Let us now have a look at what brand of set-function we *would like* to have – *positive measure* is how it is called. We start by clarifying what kind of collection of subsets (of the real line, or more generally of some space X) we want to use, in order to talk about a positive measure.

Def'n 1.5. σ-**algebra** of Subsets
Let X be a nonempty set and let $\mathcal{A}$ be a colelction of subsets of X. We say $\mathcal{A}$ is a ***σ-algebra*** of subsets of X to mean that

(a) $\varnothing \in \mathcal{A}$;

(b) for all countable $\mathcal{F} \subseteq \mathcal{A}$, $\bigcup \mathcal{F} \in \mathcal{A}$; and *closure under countable union*

(c) for all $A \in \mathcal{A}$, $X \setminus A \in \mathcal{A}$. *closure under complement*

Def'n 1.6. **Positive Measure** on a σ-algebra
Let X be a nonempty set and let $\mathcal{A}$ be a σ-algebra of subsets of X. A ***positive measure*** on $\mathcal{A}$ is a function $\mu : \mathcal{A} \to [0, \infty]$ such that

(a) $\mu(\varnothing) = 0$; and

(b) for all collection $\mathcal{F} \subseteq \mathcal{A}$ of pairwise disjoint subsets of X, $\bigcup \mathcal{F} \in \mathcal{A}$. *$\sigma$-additivity*

The framework of a positive measure on a σ-algebra is the one which works like a charm when we want to build the integration theory of Lebesgue. It will be quite desirable to streamline our considerations and make them fall in this framework!

We note that the topology $\mathcal{T}$ satisfies (a), (b) of Def'n 1.5, and the set-function $\lambda : \mathcal{T} \to [0, \infty]$ has $\lambda(\varnothing)$ and is countably additive. However, $\mathcal{T}$ is actually quite far from satisfying (c) of Def'n 1.5. We will have to look into that as we proceeds.

The remaining part of this subsection is devoted to making some easy (but useful) remarks about properties of σ-algebras and of positive measures automatically follow.

Proposition 1.4. Some Properties of a σ-algebra
Let X be a nonempty set and let $\mathcal{A}$ be a σ-algebra of X.

(a) $X \in \mathcal{A}$.

(b) For all countable $\mathcal{F} \subseteq \mathcal{A}$, $\bigcap \mathcal{F} \in \mathcal{A}$. *closure under countable intersection*

(c) For all $A, B \in \mathcal{A}$, $A \setminus B \in \mathcal{A}$. *closure under set difference.*

Proof.

(a) Since $\varnothing \in \mathcal{A}$ and $X \setminus \varnothing = X$, $X \in \mathcal{A}$.

(b) Let $\mathcal{F} \subseteq \mathcal{A}$ be countable, say $\mathcal{F} = \{F_i\}_{i \in I}$. Then

$$\bigcap \mathcal{F} = \bigcap_{i \in I} F_i = \bigcap_{i \in I} X \setminus (X \setminus F_i) = X \setminus \bigcup_{i \in I} (X \setminus F_i).$$

Since $\mathcal{A}$ is closed under complement and countable union, $\bigcup_{i \in I} (X \setminus F_i) \in \mathcal{A}$. Hence $\bigcap \mathcal{F} \in \mathcal{A}$.

(c) Since $X \in A$, $A \setminus B = A \cap (X \setminus B) \in \mathcal{A}$.

QED

Proposition 1.5. Monotonicity of Positive Measures

Let X be a nonempty set and let $\mathcal{A}$ be a σ-algebra of subsets of X. Let $\mu : \mathcal{A} \to [0, \infty]$ be a positive measure. Then μ is increasing: for all $A, B \in \mathcal{A}$ such that $A \subseteq B$, $\mu(A) \leq \mu(B)$.

Proof. Observe that

$$\mu(B) = \mu(A \cup (B \setminus A)) = \mu(A) + \mu(B \setminus A) \geq \mu(A).$$

QED

We remark that above proof, although it looks identical to the failed one given for Proposition 1.3, works since $\mathcal{A}$ is closed under set difference.

3. Back to Proposition 1.3

Subsection 1.2 was a bit of a detour from the main object of the present lecture, the set-function $\lambda : \mathcal{T} \to [0, \infty]$. A good detour, as the notion of positive measure on a σ-algebra will turn out to be of great importance for this course.

But, going back to the Proposition 1.3 we started with: since $\mathcal{T}$ is not a σ-algebra, Proposition 1.5 doe snot apply to $\lambda : \mathcal{T} \to [0, \infty]$, so we still don't have a proof that λ is increasing. We can still go ahead and prove Proposition 1.3 via a direct analysis of how open sets decompose into interval components. For clarity, we separate some parts of the argument as lemmas.

Lemma 1.6.

Let $a_1 < b_1, \ldots, a_n < b_n$ be in $\mathbb{R}$, where $a_1 < \cdots < a_n$, and suppose that the open intervals $(a_1, b_1), \ldots, (a_n, b_n)$ are pairwise disjoint. Then $a_1 < b_1 \leq a_2 < b_2 \leq \cdots \leq a_{n-1} < b_{n-1} \leq a_n < b_n$.

Proof. It suffices to examine the case $n = 2$ and show $b_1 \leq a_2$.

For contradiction, suppose that $b_1 > a_2$. Then observe that $(a_2, \min\{b_1, b_2\})$ is a nonempty interval contained in both $(a_1, b_1), (a_2, b_2)$. This contradicts the fact that $(a_1, b_1) \cap (a_2, b_2) = \varnothing$.

QED

Lemma 1.7.

Let $a, b \in \mathbb{R}$ with $a < b$ and let $A \subseteq \mathbb{R}$ be open and such that $A \subseteq (a, b)$. Then $\lambda(A) \leq b - a$.

Proof. The statement is clear for $A = \varnothing$, so we will assume $A \neq \varnothing$. We use the description of $\lambda(A)$ as a supremum. In view of the definition of a supermum, the required inequality $\lambda(A) \leq b - a$ will follow if we can verify the following claim.

- *Claim 1. Let $(a_1, b_1), \ldots, (a_n, b_n)$ be some pairwise disjoint interval components of A. Then $\sum_{j=1}^{n} (b_j - a_j) \leq b - a$.*

  Proof. Fix some intervals $(a_1, b_1), \ldots, (a_n, b_n)$ as in the claim. Reordering them if necessary, we may assume $a_1 < \cdots < a_n$. Then,

$$\begin{aligned}
\sum_{j=1}^{n} (b_j - a_j) &\leq \sum_{j=1}^{n} (b_j - a_j) + ((a_2 - b_1) + \cdots + (a_n - b_{n-1})) && \text{by Lemma 1.6} \\
&= (b_1 - a_1) + (a_2 - b_1) + (b_2 - a_2) + \cdots + (a_n - b_{n-1}) + (b_n - a_n) \\
&= b_n - a_1 && \textit{telescopic} \text{ sum} \\
&\leq b - a. && \text{since } A \subseteq (a, b) \text{ implies } a \leq a_1, b_n \leq b.
\end{aligned}$$

  This verifies the claim and thus the statement.

QED

Proof of Proposition 1.3

Let $A, B \in \mathcal{T}$ be such that $A \subseteq B$. We have to prove that $\lambda(A) \leq \lambda(B)$. If $\lambda(B) = \infty$, then the inequality to be proved is obvious; we will therefore assume $\lambda(B) < \infty$. We will also assume $B \neq \varnothing$.

We consider the (countable) decomposition $B = \bigcup_{i\in I} C_i$ of B into open intervals. From the assumption that $\lambda(B) < \infty$, it follows that every C_i is a bounded interval, say $C_i = (a_i, b_i)$ for some $a_i, b_i \in \mathbb{R}$ with $a_i < b_i$. The value $\lambda(B)$ is defined as $\sum_{i\in I}(b_i - a_i)$.

For every $i \in I$, let us denote $A_i = A \cap (a_i, b_i)$. This is a set in $\mathcal{T}$, as the intersection of two open sets is still open. It is clear, moreover, that $A_i \subseteq (a_i, b_i)$, hence Lemma 1.7 applies and gives us the upper bound

$$\lambda(A_i) \le b_i - a_i$$

for all $i \in I$.

We observe that $A_i \cap A_j = \varnothing$ for any distinct $i, j \in I$, since $A_i \subseteq (a_i, b_i)$, $A_j \subseteq (a_j, b_j)$ and $(a_i, b_i) \cap (a_j, b_j) = \varnothing$. Moreover,

$$\bigcup_{i\in I} A_i = \bigcup_{i\in I}(A \cap (a_i, b_i)) = A \cap \left(\bigcup_{i\in I}(a_i, b_i)\right) = A \cap B = A.$$

We thus find that

$$\begin{aligned}\lambda(A) &= \sum_{i\in I}\lambda(A_i)\\ &\le \sum_{i\in I}(b_i - a_i)\\ &= \lambda(B),\end{aligned}$$

and we conclude that $\lambda(A) \le \lambda(B)$, as required.

QED

4. Length of Compact Subsets of $\mathbb{R}$

We noticed there are some difficulties coming from the fact that the collection $\mathcal{T}$ of open subsets of $\mathbb{R}$ is not closed under complements. This issue is in fact quite drastic:

$$\forall A \in \mathcal{T}\,[A \ne \varnothing, A \ne \mathbb{R} \implies \mathbb{R} \setminus A \notin \mathcal{T}]. \qquad [1.7]$$

Indeed, taking complements of open sets gives the *closed* subsets of the real line. The statement in [1.7] thus says the only subsets of $\mathbb{R}$ which are *clopen* are the empty set $\varnothing$ and the total set $\mathbb{R}$. This is a statement that arises when discussing about the fact that *the real line is connected*.

We are now interested in defining notion of length for closed sets. It will be in fact easier to figure the things out if we focus on *compact* subsets of $\mathbb{R}$ (i.e. closed sets $K \subseteq \mathbb{R}$ which are also *bounded*). Some of the properties of compact sets (e.g. open covers, finite intersection property, sequential compactness,…) from PMATH 351 may come in handy in our discussions as well.

Now, how should we proceed in order to define the length $\lambda(K)$ for a compact subset $K \subseteq \mathbb{R}$? We do not have at our disposal some neat result about the structure of K, in the way we had when we discussed about open sets; but we can actually *fall back* on what we know about lengths of open sets, in the way described as follows. We start by enclosing K in a bounded open set A (which is possible since K is bounded – we can even arrange A to be an open interval, if we want). Then $A \setminus K$ is a bounded open set as well.

Exercise 1.5.

Let $K \subseteq \mathbb{R}$ be compact and let $A \subseteq \mathbb{R}$ be an open set containing K. Show that $A \setminus K$ is open.

Proof. Since K is compact, K is closed, so $\mathbb{R} \setminus K$ is open. This means $A \setminus K = A \cap (\mathbb{R} \setminus K)$ is open.

QED

The sets $K, A \setminus K$ are disjoint, and their union is A, since we are hoping that the notion of length is additive, we should then have

$$\lambda(K) + \lambda(A \setminus K) = \lambda(A) \in [0, \infty),$$

and we can turn this equation into a formula used to define the length of K:

$$\lambda(K) = \lambda(A) - \lambda(A \smallsetminus K). \tag{1.8}$$

Can this really work? In [1.8] we see an *idea* for how to proceed, but can we be really sure that the right-hand side of the equation does not depend on how A was chosen?

In [1.8] it feels bit uncomfortable that we are using the same letter λ bot hfor lengths of open sets (a well-defined notion studied in the preceding subsections) and for lengths of compact sets, a notion which is, for the moment, not rigorously defined. In order to avoid any confusions, let us make the following conventions of notation.

Notation 1.7. λ_{op}
We introduced the length-measuring map $\lambda : \mathcal{T} \to [0, \infty]$, where $\mathcal{T}$ is the collection of the open subsets of $\mathbb{R}$. Let us (at least temporarily) change the name of this map to λ_{op}, with the subscript *op* meant to remind us of open sets.

On the other hand, let us write

$$\mathcal{K} = \{K \subseteq \mathbb{R} : K \text{ is compact}\}.$$

The goal is to define a function $\lambda_{\text{cp}} : \mathcal{K} \to [0, \infty)$, where for every $K \in \mathcal{K}$ the number λ_{cp} is our notion of *length of K*. The subscript *cp* is meant to remind us that λ_{cp} is addressing lengths of compact sets.

The plan presented before is now phrased as follows: the new length-measuring map $\lambda_{\text{cp}} : \mathcal{K} \to [0, \infty)$ has to defined in such a way that we have the formula

$$\forall K \in \mathcal{K}\ \forall A \in \mathcal{T} \left[K \subseteq A \implies \lambda_{\text{cp}}(K) + \lambda_{\text{op}}(A \smallsetminus K) = \lambda_{\text{op}}(A)\right]. \tag{1.9}$$

As noticed before, we can try to use the formula [1.9] as a lever in order to actually *define* what is λ_{cp} (but some proof is required, in order to be certain that the definition makes sense).

We start with an easy observation involving finite subsets of an open set.

Lemma 1.8.
Let $A \subseteq \mathbb{R}$ be open and let $F \subseteq A$ be finite. Then $A \smallsetminus F$ is open and

$$\lambda_{\text{op}}(A \smallsetminus F) = \lambda_{\text{op}}(A).$$

Proof. We consider the following claim.

- *Claim 1. Let $x \in A$. Then $A \smallsetminus \{x\}$ is open and $\lambda_{\text{op}}(A \smallsetminus \{x\}) = \lambda_{\text{op}}(A)$.*

  Proof. We consider the decomposition

  $$A = \bigcup_{i \in I} C_i$$

  of A into interval components, and let i_0 be the index in I for which $x \in C_{i_0}$. Clearly, we have

  $$A \smallsetminus \{x\} = (C_{i_0} \smallsetminus \{x\}) \cup \bigcup_{i \in I : i \neq i_0} C_i. \tag{1.10}$$

  Direct inspection shows that $C_{i_0} \smallsetminus \{x\}$ is a union of two disjoint open intervals C' and C'', where

  $$\lambda_{\text{op}}(C') + \lambda_{\text{op}}(C'') = \lambda_{\text{op}}(C_{i_0}) \in [0, \infty]. \tag{1.11}$$

  So then [1.10] becomes

  $$A \smallsetminus \{x\} = C' \cup C'' \cup \left(\bigcup_{i \in I : i \neq i_0} C_i\right). \tag{1.12}$$

On the right-hand side of [1.12] we have a countable union of open intervals which are pairwise disjoint, and we can thus compute as follows:

$$\begin{aligned}\lambda_{\text{op}}(A \setminus \{x\}) &= \lambda_{\text{op}}\left(C'\right) + \lambda_{\text{op}}\left(C''\right) + \sum_{i \in I : i \neq i_0} \lambda_{\text{op}}(C_i) \\ &= \lambda_{\text{op}}(C_{i_0}) + \sum_{i \in I : i \neq i_0} \lambda_{\text{op}}(C_i) && \text{by [1.11]} \\ &= \lambda_{\text{op}}(A),\end{aligned}$$

where the last equality sign of this derivation we simply have the definition of $\lambda_{\text{op}}(A)$. *(Claim 1 is verified)*

We now proceed inductively on the cardinality $|F|$. The case when $|F| = 0$ is trivial, and the one when $|F| = 1$ is covered by Claim 1.

For the inductive step, we fix $k \geq 1$. Consider an open set $A \subseteq \mathbb{R}$ and a subset $F \subseteq A$ such that $|F| = k + 1$. We isolate one of the elements $x \in F$ and we write $F = F_0 \cup \{x\}$ where $|F_0| = k$. Then

$$A \setminus F = B \setminus \{x\}$$

where $B = A \setminus F_0$. Then by induction B is open with $\lambda_{\text{op}}(A) = \lambda_{\text{op}}(B)$, while Claim 1 gives us that $\lambda_{\text{op}}(B) = \lambda_{\text{op}}(B \setminus \{x\})$. Putting these things together we find

$$\lambda_{\text{op}}(A) = \lambda_{\text{op}}(B \setminus \{x\}) = \lambda_{\text{op}}(A \setminus F),$$

as required.

QED

Lemma 1.9.

Let $K \subseteq \mathbb{R}$ be compact. Let $(a', b'), (a'', b'')$ be open intervals containing K and let $D' = (a', b') \setminus K, D'' = (a'', b'') \setminus K$. Then

$$\left(b' - a'\right) - \lambda_{\text{op}}\left(D'\right) = \left(b'' - a''\right) - \lambda_{\text{op}}\left(D''\right) \in [0, \infty). \qquad [1.13]$$

Proof. The quantities indicated on the two sides of [1.13] are indeed numbers in $[0, \infty)$.

The first thought about the proof of [1.13] is that we have to distinguish some cases (can have $a' < a''$ or $a' = a''$ or $a' > a''$). This can, however, be avoided, if we go as follows: fix some real numbers a, b such that

$$a < \min\left(a', a''\right), b > \max\left(b', b''\right)$$

and look at the open interval (a, b); this contains both (a', b') and (a'', b''), hence in particular contains K. We consider the open set $D = (a, b) \setminus K$, and we will prove that either side of [1.13] is equal to $(b - a) - \lambda_{\text{op}}(K)$. This will prove, in particular, that the required inequality [1.13] is holding.

By symmetry, it is clearly sufficient to prove that one of the two sides of [1.13] is equal to $(b - a) - \lambda_{\text{op}}(K)$. Say we focus on checking that

$$\left(b' - a'\right) = \lambda_{\text{op}}\left(D'\right) = (b - a) - \lambda_{\text{op}}(D).$$

A bit of algebra shows the latter equation to be equivalent to

$$\lambda_{\text{op}}(D) = \lambda_{\text{op}}\left(D'\right) + \left(a' - a\right) + \left(b - b'\right) \qquad [1.14]$$

and thus it will suffice to verify the validity of [1.14].

But now, if we start from the fact that $a < a' < b' < b$ with $K \subseteq (a', b')$ and with $D = (a, b) \setminus K, D' = (a', b') \setminus K$, then direct inspection gives us the relation

$$D = D' \cup \left(a, a'\right] \cup \left[b', b\right),$$

which in turn implies the equality of open sets

$$D \setminus \left\{a', b'\right\} = D' \cup \left(a, a'\right) \cup \left(b', b\right). \qquad [1.15]$$

We note, moreover, that the union on the right-hand side of [1.15] involves three open sets that are pairwise disjoint. We can then write

$$\begin{aligned}
\lambda_{\text{op}}(D) &= \lambda_{\text{op}}\left(D \setminus \left\{a', b'\right\}\right) && \text{by Lemma 1.8} \\
&= \lambda_{\text{op}}\left(D' \cup \left(a, a'\right) \cup \left(b', b\right)\right) && \text{by [1.15]} \\
&= \lambda_{\text{op}}\left(D'\right) + \lambda_{\text{op}}\left(\left(a, a'\right)\right) + \lambda_{\text{op}}\left(\left(b', b\right)\right) \\
&= \lambda_{\text{op}}\left(D'\right) + \left(a' - a\right) + \left(b - b'\right),
\end{aligned}$$

which gives precisely [1.14] we had been left to prove.

QED

Def'n 1.8. **Length** of a Compact Set

Let $K \subseteq \mathbb{R}$ be compact. We define the ***length*** of K, denoted as $\lambda_{\text{cp}}(K)$, by the following procedure. Pick an open interval (a, b) that contains K, consider the open set $D = (a, b) \setminus K$, and define

$$\lambda_{\text{cp}}(K) = (b - a) - \lambda_{\text{op}}(D).^{1} \tag{1.16}$$

[1] The quantity on the right-hand side of [1.16] depends only on K due to Lemma 1.9.

The length-measuring map $\lambda_{\text{cp}} : \mathcal{K} \to [0, \infty)$ is now rigorously defined, and it was indeed obtained according to the plan – by using a special case of the formula [1.9], the case where the enclosing bounded open set A is an interval. For future use, we would like to have said formula [1.9] available in its general case; thus, we do some bootstrapping – from the special case of $A = (a, b)$ we move up to the one where A is a general open subset of $\mathbb{R}$.

We first go from the case $A = (a, b)$ to the one where A is a finite disjoint union of bounded open intervals.

Lemma 1.10.

Let $a_1 < b_1 \leq a_2 < b_2 \leq \cdots \leq a_k < b_k$ be real numbers, and consider the open set $A = \bigcup_{j=1}^{k}\left(a_j, b_j\right)$. Suppose that $K \subseteq \mathbb{R}$ is compact and contained in A. Then

$$\lambda_{\text{cp}}(K) + \lambda_{\text{op}}(A \setminus K) = \sum_{j=1}^{k}\left(b_j - a_j\right). \tag{1.17}$$

Proof. Consider the open interval (a_1, b_k). This clearly contains both A, K. Now,

$$\lambda_{\text{cp}}(K) = (b_k - a_1) - \lambda_{\text{op}}(D) \tag{1.18}$$

where $D = (a_1, b_k) \setminus K$. Moreover, since $K \subseteq A = \bigcup_{j=1}^{k}\left(a_j, b_j\right) \subseteq (a_1, b_k)$,

$$D = \left(\bigcup_{j=1}^{k-1}\left[b_j, a_{j+1}\right]\right) \cup (A \setminus K), \tag{1.19}$$

where the right-hand side of [1.19] is a disjoint union. Hence

$$\begin{aligned}
\lambda_{\text{op}}(D) &= \lambda_{\text{op}}\left(D \setminus \{b_1, a_2, b_2, \ldots, a_{k-1}, b_{k-1}, a_k\}\right) \\
&= \lambda_{\text{op}}\left(\left(\bigcup_{j=1}^{k-1}\left(b_j, a_{j+1}\right)\right) \cup (A \setminus K)\right) \\
&= \lambda_{\text{op}}(A \setminus K) + \sum_{j=1}^{k-1}\left(a_{j+1} - b_j\right).
\end{aligned}$$

Thus [1.18] becomes

$$\begin{aligned}\lambda_{\text{cp}}(K) &= (b_k - a_1) - \sum_{j=1}^{k-1}\left(a_{j+1} - b_j\right) - \lambda_{\text{op}}(A \setminus K)\\ &= (b_k - a_1) + \sum_{j=1}^{k-1}\left(b_j - a_{j+1}\right) - \lambda_{\text{op}}(A \setminus K)\\ &= \sum_{j=1}^{k}\left(b_j - a_j\right) - \lambda_{\text{op}}(A \setminus K),\end{aligned}$$

rearranging which gives

$$\lambda_{\text{cp}}(K) - \lambda_{\text{op}}(A \setminus K) = \sum_{j=1}^{k}\left(b_j - a_j\right),$$

which is what we intended to show.

QED

Proposition 1.11.
Let $K \subseteq \mathbb{R}$ be compact and let $A \subseteq \mathbb{R}$ be an open set that contains K. Then

$$\lambda_{\text{cp}}(K) + \lambda_{\text{op}}(A \setminus K) = \lambda_{\text{op}}(A) \in [0, \infty]. \tag{1.20}$$

Proof. It is convenient that we first set aside the case when the following happens: there exists an unbounded open interval contained in A. In this case, upon invoking Proposition 1.3 we find that $\lambda_{\text{op}}(A) \leq \lambda_{\text{op}}(U) = \infty$, hence that $\lambda_{\text{op}}(A) = \infty$. On the other hand, it is immediate that $U \setminus K$ must contain some unbounded open interval V; hence $A \setminus K \supseteq U \setminus K \supseteq V$, which implies that $\lambda_{\text{op}}(A \setminus K)$ is infinite as well. We conclude that in this case the formula [1.20] does indeed hold, with both its sides begin equal to ∞.

For the rest of the proof, we assume that there is no unbounded open interval contained in A> We consider the decomposition of A into inteval components

$$A = \bigcup_{i\in I} C_i, \tag{1.21}$$

and we note that every C_i is a bounded open interval.

We know that I can be finite or countably infinite. If I is finite, then we can arrange C_i's to look like in Lemma 1.10, and the equality [1.20] that is need here will follow from Lemma 1.10. We will thus assume that I is countably infinite.

Now comes the punchline: we have $K \subseteq A = \bigcup_{i\in I}$, so $\{C_i\}_{i\in I}$ is an *open cover* of the compact set K. Hence by the compactness of K, there exists $i_1, \ldots, i_k \in I$ such that $K \subseteq \bigcup_{j=1}^{k} C_{i_j}$.

Our next move is then to break A as the disjoint union $A = A' \cup A''$, where

$$A' = \bigcup_{j=1}^{k} C_{i_j}, A'' = \bigcup_{i\in I\setminus\{i_1,\ldots,i_k\}} C_i.$$

Lemma 1.10 applies in connection to the inclusion $K \subseteq A'$, and gives us that

$$\lambda_{\text{cp}}(K) + \lambda_{\text{op}}\left(A' \setminus K\right) = \lambda_{\text{op}}\left(A'\right).$$

So then we can write

$$\begin{aligned}\lambda_{\text{op}}(A) &= \lambda_{\text{op}}\left(A'\right) + \lambda_{\text{op}}\left(A''\right)\\ &= \lambda_{\text{cp}}(K) + \lambda_{\text{op}}\left(A' \setminus K\right) + \lambda_{\text{op}}\left(A''\right)\\ &= \lambda_{\text{cp}}(K) + \lambda_{\text{op}}(A \setminus K), && \text{since } K \subseteq A' \text{ and } A' \cap A'' = \varnothing\end{aligned}$$

as required.

QED

So far we have defined the notion of length for open subsets of $\mathbb{R}$ and for compact subsets of $\mathbb{R}$. These were formalized as some length-measuring maps $\lambda_{\text{op}} : \mathcal{T} \to [0, \infty]$ and $\lambda_{\text{cp}} : \mathcal{K} \to [0, \infty)$, where $\mathcal{T}, \mathcal{K}$ are the collections of open subsets of $\mathbb{R}$ and compact subsets of $\mathbb{R}$, respectively. We aim to eventually find a collection $\mathcal{M}$ of subsets of $\mathbb{R}$, called *measurable sets*, such that $\mathcal{M} \supseteq \mathcal{T} \cup \mathcal{K}$, and a length-measuring function $\lambda : \mathcal{M} \to [0, \infty]$ which extends both λ_{op} and λ_{cp}.

$\mathcal{M}$ will turn out to be a σ-algebra, and $\lambda : \mathcal{M} \to [0, \infty]$ will turn out to be a positive measure, in the sense discussed in Def'n 1.5, 1.6. This positive measure λ is the Lebesgue measure on the real line.

On our way towards constructing measurable sets, we will need to use the following fact, which is stated just in terms of open sets.

Proposition 1.12. Continuity along Decreasing Chains of Open Sets

Let $(A_n)_{n=1}^{\infty}$ be a sequence of bounded open subsets of $\mathbb{R}$ such that

$$A_1 \supseteq A_2 \supseteq \cdots \qquad [1.22]$$

and such that $\bigcap_{n=1}^{\infty} A_n = \varnothing$. Then $\lim_{n\to\infty} \lambda_{\text{op}}(A_n) = 0$.

When we arrive to examine the basic properties of the Lebesgue measure $\lambda : \mathcal{M} \to [0, \infty]$, we will have a general result about the continuity of λ along a decreasing chain. Proposition 1.12 is a special case of that result. But it would not be all right if we stated that result now and used it in order to derive Proposition 1.12 – that would create a circular argument (it would be a circular argument because we will use Proposition 1.12 in our construction of the Lebesgue measure).

This means we will have to find a way to prove Proposition 1.12 which is only using facts that we established about λ_{op} so far.

To prove Proposition 1.12, it will be convenient to go for the so-called *contrapositive* of what we are asked to prove. More precisely, instead of Proposition 1.12, we will focus on the following statement.

Proposition 1.13.

Let $(A_n)_{n=1}^{\infty}$ be a decreasing chain of bounded open subsets of $\mathbb{R}$. Suppose there exists $c > 0$ such that $\lambda_{\text{op}}(A_n) \geq c$ for every $n \geq 1$. Then $\bigcap_{n=1}^{\infty} A_n \neq \varnothing$.

Proving Proposition 1.13 will give us what we need, because Proposition 1.12 can be reduced to it. Here is the argument for the reduction.

Proof of Proposition 1.12 (assuming that Proposition 1.13 is true)

Let $(A_n)_{n=1}^{\infty}$ be a decreasing chain of bounded open subsets of $\mathbb{R}$, such that $\bigcap_{n=1}^{\infty} A_n = \varnothing$. Then for every $n \geq 1$, we have $\lambda_{\text{op}}(A_n) \geq \lambda_{\text{op}}(A_{n+1})$. Hence $\left(\lambda_{\text{op}}(A_n)\right)_{n=1}^{\infty}$ is a decreasing sequence in $[0, \infty)$, and we know that such a sequece is sure to be convergent to a limit $c \geq 0$. In order to prove Proposition 1.12, we have to show that $c = 0$.

Let us assume, for contradiction, that $c \neq 0$. This means that $c > 0$. From the general properties of a decreasing convergent sequence it follows that $\lambda_{\text{op}}(A_n)$ for every $n \geq 1$. This implies $\bigcap_{n=1}^{\infty} A_n \neq \varnothing$ by Proposition 1.13, so we have a contradiction.

So the assumption $c \neq 0$ leads to contradiction. Thus $c = 0$, as we had to prove.

QED

When working on Proposition 1.13, it will be convenient to first address the special case when the open sets A_n considered there are of an *easily tractable* kind, in the sense of the next definition.

Def'n 1.9. **Easily Tractable** Open Set

We will say that an open set $A \subseteq \mathbb{R}$ is ***easily tractable***[1] when it is bounded and only has finitely many interval components.

[1] This is a term that was concocted for specific use in our discussion (i.e. this is not a standard term).

Let us elaborate a bit the meaning of this definition. If A is an open subset of $\mathbb{R}$ which is easily tractable, then every interval component of A must be bounded (because A is so), and hense it is an open interval (a, b), with $a < b$ in $\mathbb{R}$. Let us denote the number of interval components of A by k; by suitably ordering these k intervals, we arrive to the following description for A: either $A = \varnothing$ or A can be written as

$$A = \bigcup_{j=1}^{k} (a_j, b_j)$$

with $k \geq 1$ and $a_1 < b_1 \leq \cdots \leq a_k < b_k$ in $\mathbb{R}$. This really is an *easily tractable* description.

In order to get better acquainted with easily tractable open sets, here is a little exercise.

Exercise 1.6.

Let A, B be easily tractable open subsets of $\mathbb{R}$. Prove that $A \cap B$ is an easily tractable open set as well.

Proof. Clearly $A \cap B$ is bounded.

Consider decompositions $A = \bigcup_{j=1}^{n} A_j$, $B = \bigcup_{k=1}^{m} B_k$ of A, B into interval components, respectively. Then note that

$$A \cap B = \bigcup_{j=1}^{n} \bigcup_{k=1}^{m} A_j \cap B_k,$$

which is a union of disjoint open intervals. Since there are finitely many $A_j \cap B_k$'s, $A \cap B$ is easily tractable.

QED

Consider now the following statement, which is a special case of Proposition 1.13.

Proposition 1.14.

Let $(A_n)_{n=1}^{\infty}$ be a decreasing chain of subsets of $\mathbb{R}$, where every A_n is an easily tractable open set. Suppose there exists a constant $c > 0$ such that $\lambda_{\text{op}}(A_n) \geq c$ for every $n \geq 1$. Then $\bigcap_{n=1}^{\infty} A_n \neq \varnothing$.

Our attack on Proposition 1.13 is like this:

(a) prove its special case stated in Proposition 1.14; and

(b) prove that the general case of Proposition 1.13 can be reduced to the special case from Proposition 1.14.

We first record a general compactness trick that can yield a nonempty intersection for a decreasing chain of open sets.

Lemma 1.15.

Let $(A_n)_{n=1}^{\infty}$ be a decreasing chain of subsets of $\mathbb{R}$, where every A_n is a bounded nonempty open set. Suppose that

$$\text{cl}(A_{n+1}) \subseteq A_n$$

for all $n \geq 1$. Then $\bigcap_{n=1}^{\infty} A_n \neq \varnothing$.

Proof. Let $K_n = \text{cl}(A_n)$ for all $n \geq 1$. Then each K_n is closed as a closure of a set, and bounded since A_n is bounded. Hence each K_n is compact. Now, by applying cl to the chain $A_1 \supseteq A_2 \supseteq \cdots$, we obtain

$$K_1 \supseteq K_2 \supseteq \cdots.$$

This is a nested sequence of nonempty compact sets, so by the finite intersection property of compact sets, it follows that $\bigcap_{n=1}^{\infty} K_n \neq \varnothing$. Let us then pick $x \in \bigcap_{n=1}^{\infty} K_n$, and observe that

$$x \in K_{n+1} = \text{cl}(A_{n+1}) \subseteq A_n$$

for all $n \geq 1$, which means $x \in A_n$ for all $n \geq 1$. Thus $x \in \bigcap_{n=1}^{\infty} A_n$, and this intersection is therefore a nonempty set.

QED

We would like to use Lemma 1.15 towards the proof of Proposition 1.14. But there is a problem: in the framework of Proposition 1.14, we generally do not have strong inclusions $\text{cl}(A_{n+1}) \subseteq A_n$ of the kind that were taken as hypothesis in Lemma 1.15.

For illustration, consider the simple example where every A_n is an open interval $A_n = (0, b_n)$ with $b_1 > b_2 > \cdots$, a strictly decreasing sequence of numbers in $(0, \infty)$. In order for these A_n's to satisfy the hypothesis of Proposition 1.14, we need to have

$b_n \geq c$ for some $c > 0$. This ensures that $\bigcap_{n=1}^{\infty} A_n$ contains the interval $(0, c)$, and is therefore nonempty. On the other hand, these A_n's do not satisfy the hypothesis of Lemma 1.15: for instance, $\mathrm{cl}(A_2) = [0, b_2] \nsubseteq (0, b_1) = A_1$.

In the example mentioned in the preceding paragraph, we see that the hypothesis of Lemma 1.15 would actually kick in if we would *trim* a bit the left endpoints of the A_n's. This is the idea which we will follow (and will turn out to work, once we set the things in the right way).

So then let us make a formal definition for what it means to *trim the endpoints of intervals* for an easily tractable open set.

Notation 1.10. $\mathrm{trim}_\varepsilon$

Suppose we are given $\varepsilon > 0$ and an easily tractable open set $A \subseteq \mathbb{R}$ such that $\lambda_{\mathrm{op}}(A) > \varepsilon$. We define a new open set, denoted as $\mathrm{trim}_\varepsilon(A)$, as follows: denoting the interval components of A as $(a_1, b_1), \ldots, (a_k, b_k)$, we put

$$\mathrm{trim}_\varepsilon(A) = \bigcup_{j=1: b_j - a_j > \frac{\varepsilon}{k}}^{k} \left(a_j + \frac{\varepsilon}{2k}, b_j - \frac{\varepsilon}{2k}\right). \tag{1.23}$$

In words, the trimmed set $\mathrm{trim}_\varepsilon(A)$ is obtained by distributing ε among the k component intervals $(a_1, b_1), \ldots, (a_k, b_k)$ of A, and by trying to trim each of these intervals by a length of $\frac{\varepsilon}{k}$. When doing so, for every $j \in \{1, \ldots, k\}$ we find one of two possibilities:

(a) the length $b_j - a_j$ of (a_j, b_j) is at most $\frac{\varepsilon}{k}$; the interval (a_j, b_j) is simply removed; or

(b) $b_j - a_j > \frac{\varepsilon}{k}$, in which case we shorten the interval (a_j, b_j) by removing a piece of length $\frac{\varepsilon}{2k}$ at each of its ends.

We note that the union indicated in [1.23] is always sure to have some sets in it; that is, there exists some index $j \in \{1, \ldots, k\}$ for which $b_j - a_j > \frac{\varepsilon}{k}$. Indeed, if we had $b_j - a_j \leq \frac{\varepsilon}{k}$ for all $j \in \{1, \ldots, k\}$, then summing over j in these inequalities would give a contradiction with the assumption that $\lambda_{\mathrm{op}}(A) > \varepsilon$.

The next lemma records some properties of trimmed open sets which follow directly from the definition.

Lemma 1.16.

Let $\varepsilon > 0$ and let $A \subseteq \mathbb{R}$ be an easily tractable open subset of $\mathbb{R}$ such that $\lambda_{\mathrm{op}}(A) > \varepsilon$. Then $\mathrm{trim}_\varepsilon(A)$ is an easily tractable open set with the properties that

(a) $\lambda_{\mathrm{op}}(\mathrm{trim}_\varepsilon(A)) \geq \lambda_{\mathrm{op}}(A) - \varepsilon$; and

(b) $\mathrm{cl}(\mathrm{trim}_\varepsilon(A)) \subseteq A$.

Proof. Let $A = \bigcup_{j=1}^{k} A_j$ be the decomposition of A into interval components, and assume without loss of generality that $A_1, \ldots, A_n$ have lengths more than $\frac{\varepsilon}{k}$ but $A_{n+1}, \ldots, A_k$ do not. This means

$$\mathrm{trim}_\varepsilon(A) = \bigcup_{j=1}^{n} \mathrm{trim}_{\frac{\varepsilon}{n}}(A_j). \tag{1.24}$$

From [1.24] it is immediate that $\mathrm{trim}_\varepsilon(A)$ is an easily tractable open set.

Now note that

$$\begin{aligned}\lambda_{\mathrm{op}}(\mathrm{trim}_\varepsilon(A)) &= \sum_{j=1}^{n} \lambda_{\mathrm{op}}\left(\mathrm{trim}_{\frac{\varepsilon}{n}}(A_j)\right) = \sum_{j=1}^{n}\left(\lambda_{\mathrm{op}}(A_j) - \frac{\varepsilon}{n}\right) = \left(\sum_{j=1}^{n} \lambda_{\mathrm{op}}(A_j)\right) - \varepsilon \\ &\leq \left(\sum_{j=1}^{k} \lambda_{\mathrm{op}}(A_j)\right) - \varepsilon = \lambda_{\mathrm{op}}(A) - \varepsilon,\end{aligned}$$

where the second equality is by Lemma 1.8. This verifies (a).

Also note that, given an open interval (a, b), with $a < b$, and $\eta < b - a$, we have

$$\mathrm{cl}\left(\mathrm{trim}_\eta((a, b))\right) = \left[a + \frac{\eta}{2}, b - \frac{\eta}{2}\right] \subseteq (a, b). \tag{1.25}$$

Applying [1.25] to [1.24] gives

$$\mathrm{cl}\left(\mathrm{trim}_{\varepsilon}(A)\right) = \mathrm{cl}\left(\bigcup_{j=1}^{n} \mathrm{trim}_{\frac{\varepsilon}{n}}\left(A_j\right)\right) = \bigcup_{j=1}^{n} \mathrm{cl}\left(\mathrm{trim}_{\frac{\varepsilon}{n}}\left(A_j\right)\right) \subseteq \bigcup_{j=1}^{n} A_j \subseteq \bigcup_{j=1}^{k} A_j = A.$$

This verifies (b).

QED

Proof of Proposition 1.14

We are given a decreasing chain $(A_n)_{n=1}^{\infty}$ where very A_n is an easily tractable open subset of $\mathbb{R}$, and we are given $c > 0$ such that $\lambda_{\mathrm{op}}(A - n) \geq c$ for every $n \geq 1$. Our goal is to prove that $\bigcap_{n=1}^{\infty} A_n \neq \varnothing$.

In order to reach the desired conclusion, we will construct a triangular array of sets $(A_{m,n})_{n \geq m \geq 1}$. We can think of the $A_{m,n}$'s as sitting on successive rows,as shown below, and we will construct them recursively, row by row.

$$\begin{matrix} A_{1,1} & A_{1,2} & A_{1,3} & \cdots \\ & A_{2,2} & A_{2,3} & \cdots \\ & & A_{3,3} & \cdots \\ & & & \ddots \end{matrix} \tag{1.26}$$

We will arrange the things such that, for every $m \geq 1$, the sets $A_{m,n}$'s that appear on the mth row in [1.26] have the follwing properties

(a) for all $n \geq m$, $A_{m,n}$ is an easily tractable open subset of $\mathbb{R}$;

(b) for all $n \geq m$, $\lambda_{\mathrm{op}}(A_{m,n}) \geq c\left(\frac{1}{2} + \frac{1}{2^m}\right)$; and

(c) $A_{m,m} \supseteq A_{m,m+1} \supseteq \cdots$.

So, let us describe how we do the recursive construction of rows in [1.26]. For the top row, with $m = 1$, we simply put $A_{1,n} = A_n$ for all $n \geq 1$. This clearly satisfies the conditions listed above, where (b) comes precisely to the hypothesis that $\lambda_{\mathrm{op}}(A_n) \geq c$ for every $n \geq 1$.

Now suppose that, for some $m \geq 1$, we have constructed the mth row of the array [1.26], in such a way that the listed properties are holding. We then define

$$A_{m+1,n} = A_{m,n} \cap \mathrm{trim}_{\frac{c}{2^{m+1}}}(A_{m,m}) \tag{1.27}$$

for all $n \geq m+1$. Note that the set $\mathrm{trim}_{\frac{c}{2^{m+1}}}(A_{m,m})$ mentioned in [1.27] is indeed well-defined, since we know that $\lambda_{\mathrm{op}}(A_{m,m}) \geq c\left(\frac{1}{2} + \frac{1}{2^m}\right) > \frac{c}{2^{m+1}}$. Moreover, we note that $A_{m+1,n}$ is an easily tranctable open subset of $\mathbb{R}$. This is because each of $A_{m,n}$ and $\mathrm{trim}_{\frac{c}{2^{m+1}}}(A_{m,m})$ are easily tractable, and we invoke Exercise 1.6.

We divide the remaining part of the proof into several claims.

- *Claim 1. The sequence $(A_{m+1,n})_{n=m+1}^{\infty}$ defined in [1.27] satisfies (with $m+1$ replacing m) the conditions (a), (b), (c).*

  Proof. Condition (a) was verified just before the statement of the claim. Condition (c) is immediate, since for every $n \geq m+1$, we have

$$\begin{aligned} A_{m,n+1} \subseteq A_{m,n} &\implies A_{m,n+1} \cap \mathrm{trim}_{\frac{c}{2^{m=1}}}(A_{m,m}) \subseteq A_{m,n} \cap \mathrm{trim}_{\frac{c}{2^{m=1}}}(A_{m,m}) \\ &\implies A_{m+1,n+1} \subseteq A_{m+1,n}. \end{aligned}$$

  We are thus left to verify (b), which says that $\lambda_{\mathrm{op}}(A_{m+1}, n) \geq c\left(\frac{1}{2} + \frac{1}{2^{m+1}}\right)$, for every $n \geq m+1$. In order to get this, we observe the inclusion

$$A_{m+1,n} \supseteq \mathrm{trim}_{\frac{c}{2^{m+1}}}(A_{m,n}) \tag{1.28}$$

  for all $n \geq m+1$. The inclusion [1.28] holds because $\mathrm{trim}_{\frac{c}{2^{m+1}}}(A_{m,n})$ is contained in each of the two sets that are intersected when we define $A_{m+1,n}$. That is, we have $\mathrm{trim}_{\frac{c}{2^{m+1}}}(A_{m,n}) \subseteq A_{m,n}$, and we also have $\mathrm{trim}_{\frac{c}{2^{m+1}}}(A_{m,n}) \subseteq \mathrm{trim}_{\frac{c}{2^{m+1}}}(A_{m,m})$, which is obtained by applying the $\mathrm{trim}_{\frac{c}{2^{m+1}}}$ operation in the inclusion $A_{m,n} \subseteq A_{m,m}$.

We then apply λ_{op} in the inclusion [1.28] and find that

$$\begin{aligned}\lambda_{\text{op}}(A_{m+1,n}) \geq \lambda_{\text{op}}\left(\operatorname{trim}_{\frac{c}{2^{m+1}}}(A_{m,n})\right) &\geq \lambda_{\text{op}}(A_{m,n}) - \frac{c}{2^{m+1}} \\ &\geq c\left(\frac{1}{2}+\frac{1}{2^m}\right) - \frac{c}{2^{m+1}} \geq c\left(\frac{1}{2}+\frac{1}{2^{m+1}}\right),\end{aligned}$$

exactly as we wanted. *(Claim 1 is verified)*

- *Claim 2. For every $n \geq m+1$, $A_{m+1,n} \subseteq A_{m,n}$ and $\operatorname{cl}(A_{m+1,n}) \subseteq A_{m,m}$.*

  Proof. The first inclusion is clear from [1.27]. For the second inclusion, we use the other set in the intersection indicated in [1.27], and we combine that with Lemma 1.16:

$$A_{m+1,n} \subseteq \operatorname{trim}_{\frac{c}{2^{m+1}}}(A_{m,m})$$

  implies

$$\operatorname{cl}(A_{m+1,n}) \subseteq \operatorname{cl}\left(\operatorname{trim}_{\frac{c}{2^{m=1}}}(A_{m,m})\right) \subseteq A_{m,m}.$$ *(Claim 2 is verified)*

The reward for working to construct the triangular array of $A_{m,n}$'s now comes with the following observation.

- *Claim 3. The sequence of open sets $(A_{m,m})_{m=1}^{\infty}$ satisfy the hypotheses of Lemma 1.15. That is, each $A_{m,m}$ is a bounded open set and $\operatorname{cl}(A_{m+1,m+1}) \subseteq A_{m,m}$ for all $m \geq 1$.*

  Proof. Every $A_{m,m}$ is a bounded open set (as a consequence of being easily tractable) and is nonempty, since $\lambda_{\text{op}}(A_{m,m}) \geq c\left(\frac{1}{2}+\frac{1}{2^m}\right) > 0$. Moreover, the second inclusion recorded in Claim 2 says in particular that $\operatorname{cl}(A_{m+1,m+1}) \subseteq A_{m,m}$ for every $m \geq 1$. Thus all the hypotheses of Lemma 1.15 are being satisfied. *(Claim 3 is verified)*

Lemma 1.15 gives us that $\bigcap_{m=1}^{\infty} A_{m,m} \neq \varnothing$. But for every $m \geq 1$, a repeated use of the first inclusion of Claim 2 gives

$$A_{m,m} \subseteq A_{m-1,m} \subseteq \cdots \subseteq A_{1,m} = A_m.$$

Hence if we pick $x \in \bigcap_{m=1}^{\infty} A_{m,m}$. this x will also belong to $\bigcap_{m=1}^{\infty} A_m$; the letter set is therefore nonempty, as we had to show.

QED

6. Lebesgue Measure of Bounded Measurable Sets

In this subsection we get acquainted with some of the *Lebesgue measurable* subsets of $\mathbb{R}$ (namely, those that are bounded), and with how *Lebesgue measure* is measuring such sets. More precisely, we will introduce a collection $\mathcal{M}_{\text{bdd}}$ of (some) bounded subsets of $\mathbb{R}$ and a length-measuring map $\lambda_{\text{bdd}} : \mathcal{M}_{\text{bdd}} \to [0,\infty)$, such that λ_{bdd} agrees with λ_{op} on $\mathcal{T}$ (the collection of all open subsets of $\mathbb{R}$) and λ_{bdd} agrees with λ_{cp} on $\mathcal{K}$ (the collection of all compact subsets of $\mathbb{R}$).

Def'n 1.11. **Lebesgue Measurable** Bounded Set
Let $A \subseteq \mathbb{R}$ be bounded. We say that A is ***Lebesgue measurable*** if

$$\forall \varepsilon > 0 \exists G, K \subseteq \mathbb{R}\left[G \text{ is bounded and open}, K \text{ is compact}, K \subseteq A \subseteq G, \lambda_{\text{op}}(G) - \lambda_{\text{cp}}(K) < \varepsilon\right]. \quad [1.29]$$

That is, given any $\varepsilon > 0$, we can find bounded open $G \subseteq \mathbb{R}$ and compact $K \subseteq \mathbb{R}$ such that $K \subseteq A \subseteq G$ and that $\lambda_{\text{op}}(G) - \lambda_{\text{cp}}(K) < \varepsilon$.

We shall write

$$\mathcal{M}_{\text{bdd}} = \{A \subseteq \mathbb{R} : A \text{ is bounded and Lebesgue measurable}\}.$$

Note that, in [1.29], $\lambda_{\text{op}}(G)$ is sure to be a finite quantity, since G is assuemd to be bounded. More formally speaking, there are $a, b \in \mathbb{R}$ such that $G \subseteq (a,b)$, which means $\lambda_{\text{op}}(G) = b - a < \infty$.

In fact, the last condition in [1.29], $\lambda_{\text{op}}(G) - \lambda_{\text{cp}}(K) < \varepsilon$ can be phrased in the form

$$\lambda_{\text{op}}(G \setminus K) < \varepsilon. \quad [1.30]$$

Indeed, $G \setminus K$ is open, and the definition of $\lambda_{\text{cp}}(\cdot)$ was made in such a way that we have

$$\lambda_{\text{cp}}(K) + \lambda_{\text{op}}(G \setminus K) = \lambda_{\text{op}}(G).$$

Hence the difference $\lambda_{\text{op}}(G) - \lambda_{\text{cp}}(K)$ is precisely equal to $\lambda_{\text{op}}(G \setminus K)$.

Proposition 1.17.
Let $A \in \mathcal{M}_{\text{bdd}}$. Then

$$\inf\left\{\lambda_{\text{op}}(G) : G \text{ is a bounded open set containing } A\right\} = \sup\left\{\lambda_{\text{op}}(K) : K \text{ is a compact set contained in } A\right\}. \qquad [1.31]$$

Proof. Consider the set

$$T = \left\{\lambda_{\text{cp}}(K) : K \text{ is a compact set contained in } A\right\}.$$

Clearly, T is nonempty – for instance $0 \in T$, as we find by looking at the compact set $K = \varnothing \subseteq A$. We want to argue that T is bounded above. To that end, we make the following observation.

- *Claim 1. Let G be a bounded open subset of $\mathbb{R}$ such that $G \subseteq A$ (such sets are sure to exist, due to the assumption that A is bounded). Then $\lambda_{\text{op}}(G)$ is an upper bounded for the set T.*

  Proof. We have to show that $\lambda_{\text{op}} \geq t$ for all $t \in T$. Let us fix a $t \in T$ for which we verify this inequality. We pick a compact set $K \subseteq A$ such that $\lambda_{\text{cp}}(K) = t$, and we argue like this: from $K \subseteq A \subseteq G$, it follows in particular that $K \subseteq G$. We thus find that

  $$\lambda_{\text{op}}(G) = \lambda_{\text{cp}}(K) + \lambda_{\text{op}}(G \setminus K) = t + \lambda_{\text{op}}(G \setminus K) \geq t,$$

  as required. *(Claim 1 is verified)*

Due to Claim 1, it makes sense to consider the quantity $\sup(T)$. Moreover, Claim 1 gives us that

$$\sup(T) \leq \lambda_{\text{op}}(G) \qquad [1.32]$$

for any bounded open set containing A.

Now, let us consider the set of nonnegative numbers

$$S = \left\{\lambda_{\text{op}}(G) : G \text{ is a bounded open set containing } A\right\}.$$

This is nonempty and bounded below (by 0), thus it makes sense to consider the quantity $\inf(S)$. The inequality [1.32] implies

$$\sup(T) \leq \inf(S).$$

Hence, we are left to prove the opposite inequality

$$\inf(S) \leq \sup(T). \qquad [1.33]$$

In order to establish [1.33], we will resort to the old trick of showing that

$$\forall n \in \mathbb{N}\left[\inf(S) \leq \sup(T) + \frac{1}{n}\right]. \qquad [1.34]$$

Now, we know that for every $n \in \mathbb{N}$ we can find a compact set $K_n \subseteq A$ and a bounded open set $G_n \subseteq \mathbb{R}$ containing A, such that $\lambda_{\text{op}}(G_n) - \lambda_{\text{cp}}(K_n)$. This uses the hypothesis on A which is provided in [1.29], with $\frac{1}{n}$ playing the role of ε. We thus have:

$$\begin{aligned}\inf(S) &\leq \lambda_{\text{op}}(G_n)\\ &\leq \lambda_{\text{cp}}(K_n) + \frac{1}{n}\\ &\leq \sup(T) + \frac{1}{n},\end{aligned}$$

and [1.34] follows.

QED

Def'n 1.12. **Lebesgue Measure** of a Bounded Lebesgue Measurable Set

Let $A \subseteq \mathbb{R}$ be bounded Lebesgue measurable. Then the quantity appearing in either side of [1.31] is called the ***Lebesgue measure*** of A.

In our further discussions it will come in handy to know that one can use a sequential approach to the definition of $\mathcal{M}_{\text{bdd}}$ and the map λ_{bdd}, as follows.

Proposition 1.18.

Let $A \subseteq \mathbb{R}$ be bounded. Suppose there exist a sequence $(G_n)_{n=1}^{\infty}$ of bounded open subsets of $\mathbb{R}$ and a sequence $(K_n)_{n=1}^{\infty}$ of compact subsets of $\mathbb{R}$ such that $K_n \subseteq A \subseteq G_n$ for all $n \geq 1$, and such that

$$\lim_{n\to\infty} \lambda_{\text{op}}(G_n) - \lambda_{\text{cp}}(K_n) = 0. \tag{1.35}$$

Then

(a) $A \in \mathcal{M}_{\text{bdd}}$; and

(b) both $\left(\lambda_{\text{op}}(G_n)\right)_{n=1}^{\infty}, \left(\lambda_{\text{cp}}(K_n)\right)_{n=1}^{\infty}$ are convergent, to the same limit $\lambda_{\text{bdd}}(A)$.

Proof.

(a) Suppose $\varepsilon > 0$ is given. Then we can find $n \in \mathbb{N}$ such that $\lambda_{\text{op}}(G_n) - \lambda_{\text{cp}}(K_n) < \varepsilon$ by [1.35]. This means G_n, K_n are such that G_n is bounded and open, K_n is compact, $K_n \subseteq A \subseteq G_n$, and $\lambda_{\text{op}}(G_n) - \lambda_{\text{cp}}(K_n) < \varepsilon$. Thus A is Lebesgue measurable.

(b) In [1.31] we have a description of $\lambda_{\text{bdd}}(A)$ in the form of an infimum, which implies that $\lambda_{\text{op}}(G_n) \geq \lambda_{\text{bdd}}(A)$ for all $n \geq 1$. In [1.31] we also have a description of $\lambda_{\text{bdd}}(A)$ in the form of a supremum, which implies that $\lambda_{\text{cp}}(K_n) \leq \lambda_{\text{bdd}}(A)$ for every $n \geq 1$. Upon processing a bit these inequalities, we find that

$$\forall n \in \mathbb{N}\left[\lambda_{\text{bdd}}(A) \leq \lambda_{\text{op}}(G_n) \leq \lambda_{\text{bdd}}(A) + \left(\lambda_{\text{op}}(G_n) - \lambda_{\text{cp}}(K_n)\right)\right]. \tag{1.36}$$

But our hypothesis [1.35] implies that $\lim_{n\to\infty} \lambda_{\text{bdd}}(A) + \left(\lambda_{\text{op}}(G_n) - \lambda_{\text{cp}}(K-n)\right) = \lambda_{\text{bdd}}(A)$. Hence by applying the squeeze theorem to [1.36], we obtain that $\left(\lambda_{\text{op}}(G-n)\right)_{n=1}^{\infty}$ is convergent, with limit equal to $\lambda_{\text{bdd}}(A)$.

Finally, we write

$$\forall n \in \mathbb{N}\left[\lambda_{\text{cp}}(K_n) = \lambda_{\text{op}}(G_n) - \left(\lambda_{\text{op}}(G_n) - \lambda_{\text{cp}}(K_n)\right)\right]. \tag{1.37}$$

By letting $n \to \infty$ in [1.37] we find that $\lim_{n\to\infty} \lambda_{\text{cp}}(K_n)$ exists and is equal to $\lambda_{\text{bdd}}(A)$ as well.

QED

Def'n 1.13. G_δ Set

A set $A \subseteq \mathbb{R}$ is said to be G_δ when it is a countable union of open subsets of $\mathbb{R}$.

Here are some basic properties of G_δ sets.

(a) If $A \subseteq \mathbb{R}$ is G_δ, then it can be written as $A = \bigcap_{n=1}^{\infty} G_n$ where the G_n's are open and are also assumed to form a decreasing chain $G_1 \supseteq G_2 \supseteq \cdots$. This fact has an immediate proof. Start with some arbitrary writing $A = \bigcap_{n=1}^{\infty} U_n$ where U_n's are open, and put $G_1 = U_1, G_2 = U_1 \cap U_2, \ldots$. Then the G_n's are open with $\bigcap_{n=1}^{\infty} G_n = A$, and will also form a decreasing chain.

(b) Every open $G \subseteq \mathbb{R}$ is G_δ. This is clear, as we can write $G = \bigcap_{n=1}^{\infty} G_N$ where we put $G_n = G$ for every $n \in \mathbb{N}$.

(c) Every closed $F \subseteq \mathbb{R}$ is G_δ. The standard method used to write nonempty closed $F \subseteq \mathbb{R}$ as a countable intersection of open sets $\bigcap_{n=1}^{\infty} G_n$ is by putting

$$G_n = \bigcup_{x\in F}\left(x - \frac{1}{n}, x + \frac{1}{n}\right) \tag{1.38}$$

for all $n \in \mathbb{N}$.

The following observation can be made in connection to the formula [1.38]: note that if F is bounded (hence compact), then each open set G_n provided by [1.38] are bounded as well.

Def'n 1.14. F_σ Set
A set $B \subseteq \mathbb{R}$ is said to be F_σ when it can be written as a countable union of closed subsets of $\mathbb{R}$.

Recall that a subset of $\mathbb{R}$ is F_σ if and only if its complement is G_δ. Based on this, we obtain the counterparts of the properties of G_δ sets that we listed above.

(a) If $B \subseteq \mathbb{R}$ is F_σ, then it can be written as $B = \bigcup_{n=1}^\infty F_n$ where the F_n's are closed and are also assumed and also assumed to form an increasing chain $F_1 \subseteq F_2 \subseteq \cdots$.

(b) Every closed set is F_σ.

(c) Every open set is F_σ.

Now back to our goal, of proving that the length-leasuring function λ_{bdd} on $\mathcal{M}_{\text{bdd}}$ fits with the $\lambda_{\text{op}}, \lambda_{\text{cp}}$ from the preceding lectures.

Proposition 1.19.
Let G be a bounded open subset of $\mathbb{R}$. Then $G \in \mathcal{M}_{\text{bdd}}$ and $\lambda_{\text{bdd}} = \lambda_{\text{op}}(G)$.

Proof. We know that G is F_σ, hence we can write $G = \bigcup_{n=1}^\infty F_n$ where $(F_n)_{n=1}^\infty$ is an increasing chain of closed subsets of $\mathbb{R}$. From the fact that G is bounded it follows that every F_n is bounded as well – hence F_n is a compact set. Let us also record the observation that the set-differences $(G \setminus F_n)_{n=1}^\infty$ form a decreasing chain of bounded open sets, with

$$\bigcap_{n=1}^\infty G \setminus F_n = G \setminus \left(\bigcup_{n=1}^\infty F_n\right) G \setminus G = \varnothing. \tag{1.39}$$

Quite importantly, we can invoke the continuity of λ_{op} along decreasing chains of open sets (Proposition 1.12) in connection to [1.39], so that

$$\lim_{n\to\infty} \lambda_{\text{op}}(G \setminus F_n) = 0. \tag{1.40}$$

We can then put $K_n = F_n, G_n = G$ for all $n \in \mathbb{N}$, and this will gives a sequence of compact sets $(K_n)_{n=1}^\infty$ and a sequence of bounded open sets $(G_n)_{n=1}^\infty$ such that $K_n \subseteq G \subseteq G_n$ for all $n \in \mathbb{N}$, and such that

$$\lim_{n\to\infty} \lambda_{\text{op}}(G_n) - \lambda_{\text{cp}}(K_n) = \lim_{n\to\infty} \lambda_{\text{op}}(G_n \setminus K_n) = \lim_{n\to\infty} \lambda_{\text{op}}(G \setminus F_n) = 0.$$

Hence the criterion from Proposition 1.18 can be applied, so that $G \in \mathcal{M}_{\text{bdd}}$ and that $\lambda_{\text{bdd}}(G) = \lim_{n\to\infty} \lambda_{\text{op}}(G_n)$. But $\lambda_{\text{op}}(G_n) = \lambda_{\text{op}}(G)$ for all $n \in \mathbb{N}$, hence we arrive to the desired conclusion that $\lambda_{\text{bdd}}(G) = \lambda_{\text{op}}(G)$.

QED

In the same vein as for the preceding proposition, we have the following.

Proposition 1.20.
Let $K \subseteq \mathbb{R}$ be compact. Then $K \in \mathcal{M}_{\text{bdd}}$ and $\lambda_{\text{bdd}}(K) = \lambda_{\text{cp}}(K)$.

Exercise

When looking for an example of a bounded subset of $\mathbb{R}$ which is neither open nor compact, one of the first candidates that comes to mind is $A = \mathbb{Q} \cap [0,1]$. Is this Lebesgue measurable? We can show right away that it is. In fact, the rule of thumb is that all the *natural* subsets of $\mathbb{R}$ are Lebesgue measurable; as we will see, it takes some effort to provide an example of non-measurable set.

To be precise, what happens is that $\mathbb{Q} \cap [0,1] \in \mathcal{M}_{\text{bdd}}$ and $\lambda_{\text{bdd}}(Q \cap [0,1]) = 0$. The sets of this kind are called *negligible*. One can detect them by using the following criterion.

Proposition 1.21.
Let $A \subseteq \mathbb{R}$ be bounded. Suppose there exists a sequence $(G_n)_{n=1}^\infty$ of bounded open subsets of $\mathbb{R}$ such that $A \subseteq G_n$ for all $n \in \mathbb{N}$, and such that $\lim_{n\to\infty} \lambda_{\text{op}}(G_n) = 0$. Then it follows that $A \in \mathcal{M}_{\text{bdd}}$ and $\lambda_{\text{bdd}}(A) = 0$.

Proof. This follows from Proposition 1.18 where we use the given G_n's and also consider $K_n = \varnothing$ for all $n \in \mathbb{N}$.

QED

By using Proposition 1.21, it is easy to check that $\mathbb{Q} \cap [0,1]$ is negligible. This actually happens just because we are dealing with a countable set. So instead of $\mathbb{Q} \cap [0,1]$, let us consider any bounded countable subset $A \subseteq \mathbb{R}$ in the following exercise.

Exercise 1.7.

Let $A \subseteq \mathbb{R}$ be bounded and countable.

(a) Prove that for every $\varepsilon > 0$, there is bounded open $G \subseteq \mathbb{R}$ containing A with $\lambda_{\text{op}}(G) < \varepsilon$.

(b) By using (a) and Proposition 1.21, prove that A is negligible. That is, $A \in \mathcal{M}_{\text{bdd}}$ and $\lambda_{\text{bdd}}(A) = 0$.

Proof.

(a) Since A is countable, fix an enumeration of the elements of A, say $\{a_n\}_{n\in\mathbb{N}}$. Now consider the union

$$G = \bigcup_{n\in\mathbb{N}} \left(a_n - \frac{\varepsilon}{2^{n+2}}, a_n + \frac{\varepsilon}{2^{n+2}}\right)$$

which contains A. For each $n \in \mathbb{N}$, the interval $\left(a_n - \frac{\varepsilon}{2^{n+2}}, a_n + \frac{\varepsilon}{2^{n+2}}\right)$ has length $\frac{\varepsilon}{2^{n+1}}$, so that

$$\lambda_{\text{op}}(G) \le \sum_{n\in\mathbb{N}} \frac{\varepsilon}{2^{n+1}} = \frac{\varepsilon}{2} < \varepsilon.$$

(b) For each $n \in \mathbb{N}$, let G_n be a bounded open subset of $\mathbb{R}$ containing A with $\lambda_{\text{op}}(G_n) < \frac{1}{n}$, which exists by (a). Then $\lim_{n\to\infty} \lambda_{\text{op}}(G_n) = \lim_{n\to\infty} \frac{1}{n} = 0$, so that we can invoke Proposition 1.21 to conclude that $A \in \mathcal{M}_{\text{bdd}}$ and that $\lambda_{\text{bdd}}(A) = 0$.

QED

7. Properties of $\mathcal{M}_{\text{bdd}}$ and of λ_{bdd}

We identified a collection $\mathcal{M}_{\text{bdd}}$ of *bounded measurable subsets* of $\mathbb{R}$. Moreover, we introduced a length-measuring function $\lambda_{\text{bdd}} : \mathcal{M}_{\text{bdd}} \to [0,\infty)$, and we found that

(a) every bounded open subset A of $\mathbb{R}$ belongs to $\mathcal{M}_{\text{bdd}}$ with $\lambda_{\text{bdd}}(A) = \lambda_{\text{op}}(A)$; and

(b) every compact subset K of $\mathbb{R}$ belongs to $\mathcal{M}_{\text{bdd}}$ with $\lambda_{\text{bdd}}(K) = \lambda_{\text{cp}}(K)$.

We will see that $\mathcal{M}_{\text{bdd}}$ is closed under certain set-operations. More precisely, we will establish that bounded measurable subsets of $\mathbb{R}$ forms what is called a *ring* of subsets of $\mathbb{R}$. We will also observe an additivity property of the length-measuring function λ_{bdd}>

Recall that we looked at the collection $\mathcal{T}$ of open subsets of $\mathbb{R}$, and we discovered the shortcoming that $\mathcal{T}$ is not closed under set-difference. One of the things we will find is that $\mathcal{M}_{\text{bdd}}$ is closed under set-difference:

$$\forall A_1, A_1 \in \mathcal{M}_{\text{bdd}}\,[A_1 \setminus A_2 \in \mathcal{M}_{\text{bdd}}].$$

This is a step towards resolving our issue about set-differences of open sets: if G_1, G_2 are bounded open sets, then $G_1 \setminus G_2$ may no longer be open, but it will still be in $\mathcal{M}_{\text{bdd}}$. This provides us with a notion of length for $G_1 \setminus G_2$, namely it is the number $\lambda_{\text{bdd}}(G_1 \setminus G_2)$.

Def'n 1.15. **Ring** of Sets

Let X be a nonempty set and let $\mathcal{R}$ be a collection of subsets of X. We say $\mathcal{R}$ is a ***ring*** of subsets of X to mean that

(a) $\varnothing \in \mathcal{R}$;

(b) for all $A_1, A_2 \in \mathcal{R}$, $A_1 \cup A_2 \in \mathcal{R}$; and *closure under union*

(c) for all $A_1, A_2 \in \mathcal{R}$, $A_1 \setminus A_2 \in \mathcal{R}$. *closure under set-difference*

Let X be a nonempty set and let $\mathcal{R}$ be a ring of subsets of X.

(a) $\mathcal{R}$ is closed under intersection. This is found by writing that

$$A_1 \cap A_2 = A_1 \setminus (A_1 \setminus A_2)$$

for any $A_1, A_2 \in \mathcal{R}$.

(b) By induction, $\mathcal{R}$ is closed under any finite union or intersection.

In conclusion, we can do any kind of set-operations we want, involving *finitely many* sets from $\mathbb{R}$, and we can be certain that the result of these operations will still belong to $\mathcal{R}$.

In the process of constructing the Lebesgue measure we constantly have to fall back on elementary statements about the length-measuring function λ_{op} for open sets. This happens because λ_{op} is the fundamental notion we are building on.

Exercise 1.8.

Consider a bounded open interval (a, b) with $a < b$ in $\mathbb{R}$. Suppose G_1, G_2 are open subsets of $\mathbb{R}$ such that $G_1 \cup G_2 = (a, b)$. Prove that

$$b - a \leq \lambda_{\text{op}}(G_1) + \lambda_{\text{op}}(G_2).$$

Proof. It suffices to check that

$$b' - a' \leq \lambda_{\text{op}}(G_1) + \lambda_{\text{op}}(G_2)$$

for all $a', b' \in \mathbb{R}$ such that $a < a' < b' < b$. So choose such a', b'. Now note that $[a', b'] \subseteq (a, b) = G_1 \cup G_2$. This means, if we consider the decompositions of G_1, G_2 into interval components,

$$G_1 = \bigcup_{i \in I} G_{1,i}, G_2 = \bigcup_{j \in J} G_{2,j},$$

respectively, then $\left(\bigcup_{i \in I} G_{1,i}\right) \cup \left(\bigcup_{j \in J} G_{2,j}\right)$ is (the union of) an open cover for $[a', b']$. Since $[a', b']$ is compact, we can choose a finite subcover $\{G_{1,t}\}_{t=1}^{n} \cup \{G_{2,s}\}_{s=1}^{m}$. This means

$$b' - a' = \lambda_{\text{cp}}\left([a', b']\right) \leq \lambda_{\text{op}}\left(\left(\bigcup_{t=1}^{n} G_{1,t}\right) \cup \left(\bigcup_{s=1}^{n} G_{2,s}\right)\right) \leq \left(\sum_{t=1}^{n} \lambda_{\text{op}}(G_{1,t})\right) + \left(\sum_{s=1}^{m} \lambda_{\text{op}}(G_{2,s})\right) \leq \lambda_{\text{op}}(G_1) + \lambda_{\text{op}}(G_2),$$

as required.

QED

Proposition 1.22.

Let G_1, G_2 be two bounded open subsets of $\mathbb{R}$ and consider their union $G_1 \cup G_2$, which is a bounded open set as well. Then

$$\lambda_{\text{op}}(G_1 \cup G_2) \leq \lambda_{\text{op}}(G_1) + \lambda_{\text{op}}(G_2).$$

Proof. Consider the decomposition into interval components for the open sets $G_1 \cup G_2$. Since $G_1 \cup G_2$ is bounded, this decomposition has the form

$$G_1 \cup G_2 = \bigcup_{i \in I} C_i$$

with for all $i \in I$, $C_i = (a_i, b_i)$ for some $a_i, b_i \in \mathbb{R}$. For every $i \in I$, we look at the open sets

$$G_i^{(1)} = G_1 \cap C_i, G_i^{(2)} = G_2 \cap C_i.$$

Observe that $G_i^{(1)} \cup G_i^{(2)} = (G_1 \cap C_i) \cup (G_2 \cap C_i) = (G_1 \cup G_2) \cap C_i = C_i = (a_i, b_i)$, so Exercise 1.8 applies to this situation and gives us the inequality

$$\lambda_{\text{op}}\left(G_i^{(1)}\right) + \lambda_{\text{op}}\left(G_i^{(2)}\right) \geq b_i - a_i \tag{1.41}$$

for all $i \in I$. But we know that I is countable and quantities $\lambda_{\text{op}}\left(G_i^{(1)}\right), \lambda_{\text{op}}\left(G_i^{(2)}\right), b_i - a_i$ are nonnegative, so that [1.41] implies

$$\left(\sum_{i \in I} \lambda_{\text{op}}\left(G_i^{(1)}\right)\right) + \left(\sum_{i \in I} \lambda_{\text{op}}\left(G_i^{(2)}\right)\right) \geq \sum_{i \in I}(b_i - a_i). \tag{1.42}$$

We are left to observe what is the right interpretation of the three sums over I that have appeared in [1.42]. First of all, the very definition of λ_{op} gives us that

$$\lambda_{\text{op}}(G_1 \cup G_2) = \sum_{i \in I} (b_i - a_i). \qquad [1.43]$$

Hence the right-hand side of [1.42] we have the length $\lambda_{\text{op}}(G_1 \cup G_2)$.

Let us next focus on the collection of open sets $\left\{G_i^{(1)}\right\}_{i \in I}$. We observe that these sets are pairwise disjoint, since for every distinct $i, j \in I$ we have

$$G_i^{(1)} \cap G_j^{(1)} = (G_1 \cap C_i) \cap (G_1 \cap C_j) = G_1 \cap (C_i \cap C_j) = G_1 \cap \varnothing = \varnothing.$$

So we can use on this collection Proposition 1.2, which says that

$$\sum_{i \in I} \lambda_{\text{op}}\left(G_i^{(1)}\right) = \lambda_{\text{op}}\left(\bigcup_{i \in I} G_i^{(1)}\right).$$

But

$$\bigcup_{i \in I} G_i^{(1)} = \bigcup_{i \ni I} (G_1 \cap C_i) = G_1 \cap \left(\bigcup_{i \in I} C_i\right) = G_1 \cap (G_1 \cup G_2) = G_1.$$

Hence the conclusion we draw here is that we have

$$\sum_{i \in I} \lambda_{\text{op}}\left(G_i^{(1)}\right) = \lambda_{\text{op}}(G_1). \qquad [1.44]$$

A similar calculation yields the formula

$$\sum_{i \in I} \lambda_{\text{op}}\left(G_i^{(2)}\right) = \lambda_{\text{op}}(G_2). \qquad [1.45]$$

Combining [1.42], [1.43], [1.44], [1.45] gives the desired result.

QED

Corollary 1.22.1. Subadditivity of λ_{op}

For any $k \in \mathbb{N}$ and any bounded open sets $G_1, \ldots, G_k \subseteq \mathbb{R}$, we have

$$\lambda_{\text{op}}\left(\bigcup_{j=1}^{k} G_j\right) \leq \sum_{j=1}^{k} \lambda_{\text{op}}(G_j).$$

Use Induction!

We now turn to the verification that $\mathcal{M}_{\text{bdd}}$ is a ring of subsets of $\mathbb{R}$. That is, we verify that for every $A_1, A_2 \in \mathcal{M}_{\text{bdd}}$, both $A_1 \cup A_2$ and $A_1 \setminus A_2$ are in $\mathcal{M}_{\text{bdd}}$ as well.

Lemma 1.23.

Let $K_1, \ldots, K_n \subseteq \mathbb{R}$ be compact and let $G_1, \ldots, G_n \subseteq \mathbb{R}$ be bounded and open such hat $K_j \subseteq G_j$ for all $j \in \{1, \ldots, n\}$. Consider the compact set $K = \bigcup_{j=1}^n K_j$ and the bounded open set $G = \bigcup_{j=1}^n G_j$, where $K \subseteq G$. Then

(a) $G \setminus K \subseteq \bigcup_{j=1}^n G_j \setminus K_j$; and

(b) $\lambda_{\text{op}}(G \setminus K) \leq \sum_{j=1}^n \lambda_{\text{op}}(G_j \setminus K_j)$.

Proof.

(a) We have to prove that every point $x \in G \setminus K$ belongs to the union $\bigcup_{j=1}^n G_j \setminus K_j$. So pick a point $x \in G \setminus K$. We have that $x \in G$ and $x \notin K$. The latter condition amounts to $x \notin K_1 \cup \cdots \cup K_n$, and is thus saying that $x \notin K_j$ for all $j \in \{1, \ldots, n\}$. On the other hand, since $x \in G = \bigcup_{j=1}^n G_j$, there exists an $j_0 \in \{1, \ldots, n\}$ such that $x \in G_{j_0}$. For this j_0 we find that $x \in G_{j_0}$ and $x \notin K_{j_0}$, hence that

$$x \in G_{j_0} \setminus K_{j_0} \subseteq \bigcup_{j=1}^{n} G_j \setminus K_j.$$

(b) We have that $\lambda_{\text{op}}(G \setminus K) \leq \lambda_{\text{op}}\left(\bigcup_{j=1}^n G_j \setminus K_j\right) \leq \sum_{j=1}^n \lambda_{\text{op}}(G_j \setminus K_j)$.

QED

Proposition 1.24. $\mathcal{M}_{\text{bdd}}$ Is Closed under Union

Let $A_1, A_2 \in \mathcal{M}_{\text{bdd}}$. Then $A_1 \cup A_2 \in \mathcal{M}_{\text{bdd}}$.

Proof. We need to show that for every $\varepsilon > 0$, there exists a compact set $K \subseteq \mathbb{R}$ and a bounded open set $G \subseteq \mathbb{R}$ such that

$$K \subseteq A_1 \cup A_2 \subseteq G, \lambda_{\text{op}}(G \setminus K) < \varepsilon. \qquad [1.46]$$

So fix $\varepsilon > 0$. Using the hypothesis that $A_1, A_2 \in \mathcal{M}_{\text{bdd}}$ we can find compact sets $K_1, K_2 \subseteq \mathbb{R}$ and bounded open sets $G_1, G_2 \subseteq \mathbb{R}$ such that

$$K_j \subseteq A_j \subseteq G_j, \lambda_{\text{op}}(G_j \setminus K_j) < \frac{\varepsilon}{2} \qquad [1.47]$$

for all $j \in \{1, 2\}$. Unsurprisingly, we now put $K = K_1 \cup K_2$ and $G = G_1 \cup G_2$. It is clear that K is compact and G is bounded and open, with $K \subseteq A_1 \cup A_2 \subseteq G$. Moreover, Lemma 1.23 implies that

$$\lambda_{\text{op}}(G \setminus K) \leq \lambda_{\text{op}}(G_1 \setminus K_1) + \lambda_{\text{op}}(G_2 \setminus K_2) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus [1.46] is obtained.

QED

Proposition 1.25. $\mathcal{M}_{\text{bdd}}$ Is Closed under Set-difference

Let $A_1, A_2 \in \mathcal{M}_{\text{bdd}}$. Then $A_1 \setminus A_2 \in \mathcal{M}_{\text{bdd}}$.

Proof Sketch. We follow the same plan as in the proof of Proposition 1.24. Given an $\varepsilon > 0$, we find K_1, K_2, G_1, G_2 as indicated in [1.47], and we now look at the inclusion

$$K_1 \setminus G_2 \subseteq A_1 \setminus A_2 \subseteq G_1 \setminus K_2.$$

We have thus caught $A_1 \setminus A_2$ in between the compact set $K' = K_1 \setminus G_2$ and the bounded open set $G' = G_2 \setminus K_1$. Then by playing out the analogue of the verifications form Lemma 1.23, we find that

$$G' \setminus K' \subseteq (G_1 \setminus K_1) \cup (G_2 \setminus K_2),$$

with the consequence that $\lambda_{\text{op}}(G' \setminus K') \leq \lambda_{\text{op}}(G_1 \setminus K_1) + \lambda_{\text{op}}(G_2 \setminus K_2) < \varepsilon$.

QED

Corollary 1.25.1. $\mathcal{M}_{\text{bdd}}$ Is a Ring of Subsets of $\mathbb{R}$

$\mathcal{M}_{\text{bdd}}$ is a ring of subsets of $\mathbb{R}$.

Do We Need a Proof?

We now put into evidence an addition to Proposition 1.24, which kicks in if we make the extra assumption that the sets $A_1, A_2 \in \mathcal{M}_{\text{bdd}}$ we are dealing with are disjoint. Here again there is an elementary statement to fall back one – about lengths of compact sets, this time.

Exercise 1.9.

Let K_1, K_2 be compact subsets of $\mathbb{R}$ such that $K_1 \cap K_2 = \varnothing$. Consider the union $K_1 \cup K_2$, which is a compact set as well. Prove that $\lambda_{\text{cp}}(K_1 \cup K_2) = \lambda_{\text{cp}}(K_1) + \lambda_{\text{cp}}(K_2)$.

Proof. We can choose an interval (a, b) that contains both K_1, K_2, so that we may write

$$\begin{aligned} \lambda_{\text{cp}}(K_1) &= (b-a) - \lambda_{\text{op}}((a,b) \setminus K_1) \\ \lambda_{\text{cp}}(K_1 \cup K_2) &= (b-a) - \lambda_{\text{op}}((a,b) \setminus (K_1 \cup K_2)) \end{aligned}.$$

We remark that the above discussion makes sense due to Lemma 1.9. But note that, this means

$$\begin{aligned} \lambda_{\text{cp}}(K_1 \cup K_2) &= (b-a) - \lambda_{\text{op}}((a,b) \setminus (K_1 \cup K_2)) = (b-a) - \lambda_{\text{op}}(((a,b) \setminus K_1) \setminus K_2) \\ &= (b-a) - \left(\lambda_{\text{op}}((a,b) \setminus K_1) - \lambda_{\text{cp}}(K_2)\right) = \left((b-a) - \lambda_{\text{op}}((a,b) \setminus K_1)\right) + \lambda_{\text{cp}}(K_2) = \lambda_{\text{cp}}(K_1) + \lambda_{\text{cp}}(K_2). \end{aligned}$$

QED

Proposition 1.26.
Let $A_1, A_2 \in \mathcal{M}_{\text{bdd}}$ be disjoint. Then

$$\lambda_{\text{bdd}}(A_1 \cup A_2) = \lambda_{\text{bdd}}(A_1) + \lambda_{\text{bdd}}(A_2).$$

Proof. We use the sequential approach to λ_{bdd} stated in Proposition 1.18.

For every $n \in \mathbb{N}$, let us consider some compact sets $K_n^{(1)}, K_n^{(2)}$ and bounded open sets $G_n^{(1)}, G_n^{(2)}$ such that

$$K_n^{(j)} \subseteq A \subseteq G_n^{(j)}, \lambda_{\text{op}}\left(G_n^{(j)} \setminus K_n^{(j)}\right) < \frac{1}{n}$$

for all $j \in \{1, 2\}$. Then by Proposition 1.18,

$$\lim_{n\to\infty} \lambda_{\text{cp}}\left(K_n^{(j)}\right) = \lambda_{\text{bdd}}\left(A_j\right) \tag{1.48}$$

for all $j \in \{1, 2\}$. Proceeding exactly as in the proof of Proposition 1.24, we find that for all $n \in \mathbb{N}$, we have

$$K_n \subseteq A_1 \cup A_2 \subseteq G_n, \lambda_{\text{op}}(G_n \setminus K_n) < \frac{2}{n},$$

where $K_n = K_n^{(1)} \cup K_n^{(2)}$ and $G_n = G_n^{(1)} \cup G_n^{(2)}$. By invoking Proposition 1.18 again,

$$\lim_{n\to\infty} \lambda_{\text{cp}}(K_n) = \lambda_{\text{bdd}}(A_1 \cup A_2). \tag{1.49}$$

Now by using the hypothesis $A_1 \cap A_2 = \varnothing$, we note that for all $n \in \mathbb{N}$, $K_n^{(1)} \cap K_n^{(2)} = \varnothing$ as well, so that

$$\lambda_{\text{cp}}\left(K_n^{(1)} \cup K_n^{(2)}\right) = \lambda_{\text{cp}}\left(K_n^{(1)}\right) = \lambda_{\text{cp}}\left(K_n^{(2)}\right)$$

by using Exercise 1.9. This implies that

$$\lim_{n\to\infty} \lambda_{\text{cp}}(K_n) = \lim_{n\to\infty} \lambda_{\text{cp}}\left(K_n^{(1)} \cup K_n^{(2)}\right) = \lim_{n\to\infty} \lambda_{\text{cp}}\left(K_n^{(1)}\right) + \lambda_{\text{cp}}\left(K_n^{(2)}\right) = \lambda_{\text{bdd}}(A_1) + \lambda_{\text{bdd}}(A_2) \tag{1.50}$$

by using [1.48]. Combining [1.48], [1.50] gives the desired equality.

QED

8. Measurable Subsets of $\mathbb{R}$

We now take the step of introducing the full σ-algebra $\mathcal{M}$ of (possibly unbounded) Lebesgue measurable subsets of $\mathbb{R}$. Recall that we examined the length-measuring function $\lambda_{\text{op}} : \mathcal{T} \to [0, \infty]$, and we found it to have some good properties. But, nevertheless, $\mathcal{T}$ is not a σ-algebra, hence λ_{op} is not a *positive measure* used in Lebesgue integration theory. We are now about to get the fix for that: $\mathcal{T}$ will be included in $\mathcal{M}$, which is a σ-algebra, and λ_{op} will be extended to a positive measure $\lambda : \mathcal{M} \to [0, \infty]$.

We introduce the Lebesgue measure $\lambda : \mathcal{M} \to [0, \infty]$ as an upgrade of the length-measuring function $\lambda_{\text{bdd}} : \mathcal{M}_{\text{bdd}} \to [0, \infty)$ studied during the preceding couple of subsections. More precisely, we go as follows.

Def'n 1.16. **Lebesgue Measurable** Set, **Lebesgue Measure** of a Lebesgue Measurable Set
Let $M \subseteq \mathbb{R}$. If $M \cap A \in \mathcal{M}_{\text{bdd}}$ for all $A \in \mathcal{M}_{\text{bdd}}$, then we say M is ***Lebesgue measurable***. We shall denote the collection of Lebesgue measurable sets by $\mathcal{M}$.

For a Lebesgue measurable set $M \in \mathcal{M}$, we define

$$\lambda(M) = \sup_{A \in \mathcal{M}_{\text{bdd}}} \lambda_{\text{bdd}}(M \cap A),$$

which we call the ***Lebesgue measure*** of M.

The first thing we check is that the map $\lambda : \mathcal{M} \to [0, \infty]$ introduced in Def'n 1.16 extends of the map $\lambda_{\text{bdd}} : \mathcal{M}_{\text{bdd}} \to [0, \infty)$.

Proposition 1.27.
$\mathcal{M} \supseteq \mathcal{M}_{\mathrm{bdd}}$.

Proof. Given $M \in \mathcal{M}_{\mathrm{bdd}}$, $M \cap A \in \mathcal{M}_{\mathrm{bdd}}$ for all $A \in \mathcal{M}_{\mathrm{bdd}}$ since $\mathcal{M}_{\mathrm{bdd}}$ is a ring of sets.

QED

Proposition 1.28.
For all $M \in \mathcal{M}_{\mathrm{bdd}}$, $\lambda(M) = \lambda_{\mathrm{bdd}}(M)$.

Proof. Let $M \in \mathcal{M}_{\mathrm{bdd}}$. We have to check that

$$\sup_{A \in \mathcal{M}_{\mathrm{bdd}}} \lambda_{\mathrm{bdd}}(M \cap A) = \lambda_{\mathrm{bdd}}(M). \tag{1.51}$$

We prove [1.51] by double inequality.

In view of the definition of a supremum as the least upper bound, we check that $\lambda_{\mathrm{bdd}}(M)$ is an upper bound for $\lambda_{\mathrm{bdd}}(M \cap A)$ where A runs in $\mathcal{M}_{\mathrm{bdd}}$. So let us choose $A \in \mathcal{M}_{\mathrm{bdd}}$, for which we verify $\lambda_{\mathrm{bdd}}(M \cap A) \leq \lambda_{\mathrm{bdd}}(M)$. We look at the standard decomposition

$$M = (M \cap A) \cup (M \setminus A)$$

where $M \cap A$ and $M \setminus A$ are disjoint. Since $M, A \in \mathcal{M}_{\mathrm{bdd}}$ and $\mathcal{M}_{\mathrm{bdd}}$ is a ring of sets, $M \cap A, M \setminus A \in \mathcal{M}_{\mathrm{bdd}}$. This means that

$$\lambda_{\mathrm{bdd}}(M \cap A) + \lambda_{\mathrm{bdd}}(M \setminus A) = \lambda_{\mathrm{bdd}}(M),$$

from which the inequality $\lambda_{\mathrm{bdd}}(M \cap A) \leq \lambda_{\mathrm{bdd}}(M)$ follows.

On the other hand, by noting that $M \in \mathcal{M}_{\mathrm{bdd}}$, we have that $\lambda_{\mathrm{bdd}}(M) = \lambda_{\mathrm{bdd}}(M \cap M) \leq \sup_{A \in \mathcal{M}_{\mathrm{bdd}}} \lambda_{\mathrm{bdd}}(M \cap A)$. This proves the other inequality.

Thus [1.51] is established.

QED

Corollary 1.28.1.
Let $K \subseteq \mathbb{R}$ be compact. Then $K \in \mathcal{M}$ and $\lambda(K) = \lambda_{\mathrm{cp}}(K)$.

See Proposition 1.27, 1.28

Exercise 1.10.
Let $A, B \in \mathcal{M}_{\mathrm{bdd}}$ be such that $A \subseteq B$. Prove that $\lambda_{\mathrm{bdd}}(A) \leq \lambda_{\mathrm{bdd}}(B)$.

Proof. Note that

$$\lambda_{\mathrm{bdd}}(B) = \lambda_{\mathrm{bdd}}((B \setminus A) \cup (B \cap A)) = \lambda_{\mathrm{bdd}}(B \setminus A) + \lambda_{\mathrm{bdd}}(B \cap A) = \lambda_{\mathrm{bdd}}(B \setminus A) + \lambda_{\mathrm{bdd}}(A),$$

from where the desired inequality follows.

QED

For a *bounded* open set G, there is no problem to get the analogue of Corollary 1.28.1. We say that G belongs to $\mathcal{M}_{\mathrm{bdd}}$ and has $\lambda_{\mathrm{bdd}}(G) = \lambda_{\mathrm{op}}(G)$, thus Propositoin 1.28 assures us that $G \in \mathcal{M}$ and $\lambda(G) = \lambda_{\mathrm{op}}$. But we don't want to only look at bounded open sets – the unbounded open sets would better be Lebesgue measurable too.

Here is a criterion that simplifies a bit the verification of the condition from tne definition of Lebesgue measurability.

Proposition 1.29.
Let $M \subseteq \mathbb{R}$ such that

$$\forall n \in \mathbb{N}\left[M \cap (-n, n) \in \mathcal{M}_{\mathrm{bdd}}\right]. \tag{1.52}$$

Then $M \in \mathcal{M}$, and its Lebesgue measure can be obtained as

$$\lambda(M) = \lim_{n \to \infty} \lambda_{\mathrm{bdd}}(M \cap (-n, n)). \tag{1.53}$$

Proof. We divide the argument into several claims.

- *Claim 1. For every $A \in \mathcal{M}_{\text{bdd}}$, $M \cap A \in \mathcal{M}_{\text{bdd}}$.*

  Proof. Let $A \in \mathcal{M}_{\text{bdd}}$. Using the fact that A is bounded, we pick $n \in \mathbb{N}$ such that $A \subseteq (-n, n)$. Then $A \cap (-n, n) = A$, hence we can write
  $$M \cap A = M \cap ((-n, n) \cap A) = (M \cap (-n, n)) \cap A.$$
  In this way, $M \cap A$ gets to be written as the intersection of two sets that are known to be in $\mathcal{M}_{\text{bdd}}$, namely A and $M \cap (-n, n)$. Since $\mathcal{M}_{\text{bdd}}$ is closed under finite intersections, we conclude that $M \cap A \in \mathcal{M}_{\text{bdd}}$, as required.

- *Claim 2. The sequence $(\lambda_{\text{bdd}}(M \cap (-n, n)))_{n=1}^{\infty}$ is an increasing sequence of numbers in $[0, \infty)$.*

  Proof. The statement to be checked here is that
  $$\forall n \in \mathbb{N}\left[\lambda_{\text{bdd}}(M \cap (-n, n)) \leq \lambda_{\text{bdd}}(M \cap (-n-1, n+1))\right].$$
  This holds simply because $M \cap (-n, n) \subseteq M \cap (-n-1, n+1)$, and by invoking Exercise 1.10.

It is an immediate consequence of Claim 2 that $(\lambda_{\text{bdd}}(M \cap (-n, n)))_{n=1}^{\infty}$ has a limit $l \in [0, \infty]$. We now show that $l = \lambda(M)$ by using double inequalities.

- *Claim 3. $l \leq \lambda(M)$.*

  Proof. In view of how l is defined, it suffices to show that
  $$\forall n \in \mathbb{N}\left[\lambda_{\text{bdd}}(M \cap (-n, n)) \leq \lambda(M)\right]. \tag{1.54}$$
  The definition of $\lambda(M)$ in Def'n 1.16 ensures that $\lambda(M) \leq \lambda_{\text{bdd}}(M \cap A)$ for all $A \in \mathcal{M}_{\text{bdd}}$. But on the left-hand side of [1.54] we have precisely a quantity $\lambda_{\text{bdd}}(M \cap A)$ with $A = (-n, n)$.

- *Claim 4. $l \geq \lambda(M)$.*

  Proof. In view of the definition of $\lambda(M)$ as a supremum (i.e. *least* upper bound), it suffices to check that l is an upper bound for those quantities. Or more precisely, it suffices to check that
  $$\forall A \in \mathcal{M}_{\text{bdd}}\left[l \geq \lambda_{\text{bdd}}(M \cap A)\right]. \tag{1.55}$$
  So let us fix $A \in \mathcal{M}_{\text{bdd}}$, for which we will verify that [1.55] holds. Since A is bounded, there is $n \in \mathbb{N}$ such that $A \subseteq (-n, n)$. So then we have
  $$l \geq \lambda_{\text{bdd}}(M \cap (-n, n)) \geq \lambda_{\text{bdd}}(M \cap A),$$
  where for the first inequality we use the fact that the limit of an increasing sequence is at least every term of the sequence and for the second inequality we invoke Exercise 1.10, in connection to the inclusion $M \cap A \subseteq M \cap (-n, n)$.

Combining Claim 2, 3, 4 gives [1.53].

QED

The criterion provided by Propostiion 1.29 is very suitable for arguing that every open set G is Lebesgue measurable, and that $\lambda(G) = \lambda_{\text{op}}(G)$. But here there is a background property of λ_{op} that has to be mentioned. We add this to our list of such properties, and record it in the next exercise.

Exercise 1.11.

Let $(G_n)_{n=1}^{\infty}$ be an increasing chain of open subsets of $\mathbb{R}$ and consider their union $G = \bigcup_{n=1}^{\infty} G_n$. Based on the facts about λ_{op} in Subsection 1, 2, 3, prove that $\lim_{n\to\infty} \lambda_{\text{op}}(G_n) = \lambda_{\text{op}}(G)$.

tl;dr

Armed with Exercise 1.11, here is then our desired statement about open sets.

Proposition 1.30.

Let G be an open subset of $\mathbb{R}$. Then $G \in \mathcal{M}$ and $\lambda(G) = \lambda_{\text{op}}(G)$.

Proof. For all $n \in \mathbb{N}$, we have that $G \cap (-n, n)$ is a bounded open set, and in particular it belongs to $\mathcal{M}_{\text{bdd}}$. The hypothesis of Proposition 1.29 is thus satisfied, and the said proposition gives us that $G \in \mathcal{M}$ and that

$$\lambda(G) = \lim_{n\to\infty} \lambda_{\text{bdd}}(G \cap (-n, n)). \tag{1.56}$$

We now move to proving the equality $\lambda(G) = \lambda_{\text{op}}(G)$. Since λ_{bdd} extends λ_{op} on bounded open sets, we have $\lambda_{\text{bdd}}(G \cap (-n, n)) = \lambda_{\text{op}}(G \cap (-n, n))$ for all $n \in \mathbb{N}$. Hence [1.56] can also be written in the form

$$\lambda(G) = \lim_{n\to\infty} \lambda_{\text{op}}(G \cap (-n, n)). \tag{1.57}$$

But the limit on the right-hand side of [1.57] is equal to $\lambda_{\text{op}}(G)$, as we see by applying Exercise 1.11 to the increasing chain of open sets

$$G \cap (-1, 1) \subseteq G(-2, 2) \subseteq \cdots$$

which has $\bigcup_{n=1}^{\infty}(G \cap (-n, n)) = G \cap \left(\bigcup_{n=1}^{\infty}(-n, n)\right) = G \cap \mathbb{R} = G$. Thus $\lambda(G) = \lambda_{\text{op}}(G)$, as required.

QED

We now put into evidence the fact that $\mathcal{M}$ is an *algebra* of subsets of $\mathbb{R}$.

Def'n 1.17. **Algebra** of Sets

Let X be a set. We say a collection $\mathcal{A}$ of subsets of X is an ***algebra*** of subsets of X if $X \in \mathcal{A}$ and $\mathcal{A}$ is a ring of subsets of X.

This comes a bit short of the declared goal of proving that $\mathcal{M}$ is a σ-algebra of subsets of $\mathbb{R}$ – we wil lfix this shortcoming on the next subsection.

Proposition 1.31.

$\mathcal{M}$ is an algebra of subsets of $\mathbb{R}$.

Proof.

(a) Since $\varnothing, \mathbb{R}$ are open, $\varnothing, \mathbb{R} \in \mathcal{M}$.

(b) Let $M \in \mathcal{M}$. Given any $A \in \mathcal{M}_{\text{bdd}}$, we can write $(\mathbb{R} \setminus M) \cap A = A \setminus (M \cap A)$. The set $A \setminus (M \cap A)$ is in λ_{bdd}, since $\mathcal{M}_{\text{bdd}}$ is a ring of sets.

We thus found that $(\mathbb{R} \setminus M) \cap A \in \mathcal{M}_{\text{bdd}}$ for all $A \in \mathcal{M}_{\text{bdd}}$. This means that $\mathbb{R} \setminus M \in \mathcal{M}$.

In particular, given any $N \in \mathcal{M}$, $M \setminus N = (\mathbb{R} \setminus N) \cap M \in \mathcal{M}$ by (d).

(c) Let $M_1, M_2 \in \mathcal{M}$. Then for all $A \in \mathcal{M}_{\text{bdd}}$,

$$(M_1 \cup M_2) \cap A = (M_1 \cap A) \cup (M_2 \cap A) \in \mathcal{M}_{\text{bdd}}.$$

This means $M_1 \cup M_2 \in \mathcal{M}$.

(d) Let $M_1, M_2 \in \mathcal{M}$. Then for all $A \in \mathcal{M}_{\text{bdd}}$,

$$(M_1 \cap M_2) \cap A = (M_1 \cap A) \cap (M_2 \cap A) \in \mathcal{M}_{\text{bdd}}.$$

This means $M_1 \cap M_2 \in \mathcal{M}$.

Since $\mathcal{M}$ is a collection of subsets of $\mathbb{R}$ that has $\varnothing, \mathbb{R}$ and is closed under union and set-difference, $\mathcal{M}$ is an algebra of subsets of $\mathbb{R}$.

QED

We noteiced that every compact subset of $\mathbb{R}$ belongs to $\mathcal{M}$ (Corollary 1.28.1), but we did not say anything about general *closed* subsets of $\mathbb{R}$, which may be unbounded. At this point it has, however, become clear that every closed set $F \subseteq \mathbb{R}$ belongs to $\mathcal{M}$. Indeed, the complement $G = \mathbb{R} \setminus F$ is an open set, so we argue like this:

$$\begin{aligned} & G \in \mathcal{M} && \text{by Proposition 1.30} \\ \implies & \mathbb{R} \setminus G \in \mathcal{M} && \text{by Proposition 1.31} \\ \implies & F \in \mathcal{M}\,. \end{aligned}$$

Let also record the fact that the additivity property we had obtained for λ_{bdd} in Proposition 1.26 can be upgraded to the framework of λ on $\mathcal{M}$.

Proposition 1.32.
Let $M_1, M_2 \in \mathcal{M}$ be disjoint. Consider the union $M_1 \cup M_2$, which is still in $\mathcal{M}$ by Proposition 1.31. Then

$$\lambda\left(M_1 \cup M_2\right) = \lambda\left(M_1\right) + \lambda\left(M_2\right). \tag{1.58}$$

Proof. This is an easy application of the limit trick observed in Proposition 1.29, combined with the additivity property we already know for λ_{bdd}. Indeed,

$$\forall n \in \mathbb{N}\left[\left(M_1 \cup M_2\right) \cap (-n, n) = \left(M_1 \cap (-n, n)\right) \cup \left(M_2 \cap (-n, n)\right)\right], \tag{1.59}$$

where $M_1 \cap (-n, n)$ and $M_2 \cap (-n, n)$ are two pairwise disjoint sets from $\mathcal{M}_{\text{bdd}}$. Upon applying Proposition 1.26 to this situation, we find that

$$\forall n \in \mathbb{N}\left[\lambda_{\text{bdd}}\left(M_1 \cup M_2\right) \cap (-n, n) = \lambda_{\text{bdd}}\left(M_1 \cap (-n, n)\right) + \lambda_{\text{bdd}}\left(M_2 \cap (-n, n)\right)\right]. \tag{1.60}$$

When we make $n \to \infty$ in [1.60], Proposition 1.29 implies [1.50], the desired equality.

QED

In order to reach our declared goal of proving that $\mathcal{M}$ is a σ-algebra and that $\lambda : \mathcal{M} \to [0, \infty]$ is a positive measure, we have to strengthen the results of Proposition 1.31, 1.32 so that they cover *countable* unions, which is the topic of the next subsection.

9. λ Is a Positive Measure

In the preceeding subsections we introduced the collection $\mathcal{M}$ of Lebesgue measurable subsets of $\mathbb{R}$, and for every $M \in \mathcal{M}$ we defined its Lebesgue measure $\lambda(M)$ as the supremum of the lengths of all its bounded truncations:

$$\lambda(M) = \sup\left\{\lambda_{\text{bdd}}(M \cap A) : A \in \mathcal{M}_{\text{bdd}}\right\} \in [0, \infty].$$

We saw that $\lambda(M)$ can in fact be also obtained as a plain limit:

$$\lambda(M) = \lim_{n \to \infty} \lambda_{\text{bdd}}(M \cap (-n, n)).$$

We also noted that $\mathcal{M}$ is an algebra of subsets of $\mathbb{R}$, and that λ is an additive function on M. This has the following consequence.

Proposition 1.33.
Let $M, N \in \mathcal{M}$ be such that $M \subseteq N$. Then $\lambda(M) \le \lambda(N)$.

Proof. Let $M' = N \setminus M$. Then $M' \in \mathcal{M}$ since $\mathcal{M}$ is closed under intersection, so that the additivity of λ implies

$$\lambda(N) = \lambda\left(M'\right) + \lambda(M).$$

Then the fact that λ is nonnegative implies $\lambda(N) \ge \lambda(M)$, as required.

QED

The following are what we want to show in this subsection.

(a) We want to prove that $\mathcal{M}$ is closed under countable union. This will allow us to conclude that $\mathcal{M}$ is a σ-algebra.

(b) We want to prove that if $\mathcal{C} \subseteq \mathcal{M}$ is a countable collection of pairwise disjoint Lebesgue measurable sets, then

$$\lambda\left(\bigcup \mathcal{C}\right) = \sum_{M \in \mathcal{C}} \lambda(M).$$

This will allow us to conclude that λ is a positive measure.

It turns out that the key towards obtaining the above needed properties (a) and (b) is actually lying at the preceding level of our development, where we looked at the length-measuring function $\lambda_{\text{bdd}} : \mathcal{M}_{\text{bdd}} \to [0, \infty)$. So we first consider some *preparations* for proving (a) and (b), which go at the level of $\mathcal{M}_{\text{bdd}}$.

Note that λ_{bdd} definitely is not closed under infinitely countable union, since we cannot control the boundedness of such a union. However, we will see that good properties can be obtained when we *force* the boundedenss of the union, by putting it among our hypotheses. We start by making precise the setting that we want to use.

Throughout this subsection we fix a countable collection $\{A_n\}_{n=1}^{\infty}$ of pairwise disjoint sets from $\mathcal{M}_{\text{bdd}}$. We assume, moreover, there exists $r > 0$ such that $A_n \subseteq (-r, r)$ for all $n \geq 1$.

We consider the union $A = \bigcup_{n=1}^{\infty} A_n$. It is clear that $A \subseteq (-r, r)$; thus it is a bounded set, but it is not so clear whether $A \in \mathcal{M}_{\text{bdd}}$. The goal is to verify that $A \in \mathcal{M}_{\text{bdd}}$, and moreover we have the *correct* value for $\lambda_{\text{bdd}}(A)$. The key point is provided by the following lemma.

Lemma 1.34.

Let $\varepsilon > 0$. Then there is $p_0 \in \mathbb{N}$ such that there are compact sets $K_1, \ldots, K_{p_0}$ and an open set U such that

(a) $K_1 \subseteq A_1, \ldots, K_{p_0} \subseteq A_{p_0}$;

(b) $A \subseteq U \subseteq (-r, r)$; and

(c) $\lambda_{\text{op}}(U) < \varepsilon + \lambda_{\text{cp}}\left(\bigcup_{j=1}^{p_0} K_j\right)$.

Proof. Every individual set A_n can be approximated in the way indicated in the definition of $\mathcal{M}_{\text{bdd}}$ (Def'n 1.11): given any $\varepsilon' > 0$, there are a bounded open set G_n containing A_n and a compact set K_n contained A_n, with

$$\lambda_{\text{op}}(G_n) - \lambda_{\text{cp}}(K_n) < \varepsilon'.$$

We write down a sequence of such approximations, where for each of them we use a suitable fraction ε' of the ε given in the statement. That is, for every $n \geq 1$, we find a compact set K_n and a bounded open set G_n such that

$$K_n \subseteq A_n \subseteq G_n, \lambda_{\text{op}}(G_n) < \lambda_{\text{cp}}(K_n) + \frac{\varepsilon}{2^{n+1}}. \tag{1.61}$$

By replacing G_n with $G_n \cap (-r, r)$ (which is still a bounded open set containing A_n, and has $\lambda(G_n \cap (-r, r)) \leq \lambda(G_n)$) we may assume that $G_n \subseteq (-r, r)$ for all $n \geq 1$.

We then split the rest of the proof into few claims.

- *Claim 1. Let $U = \bigcup_{n=1}^{\infty} G_n$. Then U is an open subset of $\mathbb{R}$ such that $A \subseteq U \subseteq (-r, r)$.*

  Proof. U is open as a union of open sets and $U \subseteq (-r, r)$ since each G_n is a subset of $(-r, r)$. The inclusion $A \subseteq U$ is obtained by noting that $A_n \subseteq G_n$ for all $n \in \mathbb{N}$.

- *Claim 2. For every $p \in \mathbb{N}$, consider the open set $U_p = \bigcup_{n=1}^{p} G_n \subseteq U$. Then there is $p_0 \in \mathbb{N}$ such that*

$$\lambda_{\text{op}}\left(U_{p_0}\right) > \lambda_{\text{op}}(U) - \frac{\varepsilon}{2}.$$

Proof. It is immediate that $U_1 \subseteq U_2 \subseteq \cdots$ and that $\bigcup_{n=1}^{\infty} U_n = U$. We are dealing with an increasing chain of open sets, so by Exercise 1.11 we have

$$\lim_{n\to\infty} \lambda_{\text{op}}(U_n) = \lambda_{\text{op}}(U). \tag{1.62}$$

In [1.62], the limit $\lambda_{\text{op}}(U)$ is finite – indeed, from $U \subseteq (-r, r)$ it follows that $\lambda_{\text{op}}(U) \leq 2r$. So then by the definition of limits there is a $p_0 \in \mathbb{N}$ such that the claim holds.

It remains to show that the compact sets $K_1, \ldots, K_{p_0}$ and the open set U satisfy the conditions listed in the statement. Note that (a), (b) are satisfied due to how $K_1, \ldots, K_{p_0}$ and U were constructed. We are left to verify that our *splitting* of ε was made judiciously enough in order to imply that (c) holds.

Consider Lemma 1.23, which was used in the proof that $\mathcal{M}_{\text{bdd}}$ is a ring of sets. We apply this lemma in connection to the inclusions $K_1 \subseteq G_1, \ldots, K_{p_0} \subseteq G_{p_0}$ to find that

$$\lambda_{\text{op}}\left(\left(\bigcup_{n=1}^{p_0} G_n\right) \setminus \left(\bigcup_{n=1}^{p_0} K_n\right)\right) \leq \sum_{n=1}^{p_0} \lambda_{\text{op}}(G_n \setminus K_n). \tag{1.63}$$

On the right-hand side of [1.63], every $\lambda_{\text{op}}(G_n \setminus K_n)$ can be replaced with $\lambda_{\text{op}}(G_n) - \lambda_{\text{cp}}(K_n)$, and is therefore less than $\frac{\varepsilon}{2^{n+1}}$, deu to how K_n and G_n were chosen at the beginning of the proof. So then [1.63] can be continued with

$$\lambda_{\text{op}}\left(\left(\bigcup_{n=1}^{p_0} G_n\right) \setminus \left(\bigcup_{n=1}^{p_0} K_n\right)\right) \leq \sum_{n=1}^{p_0} \frac{\varepsilon}{2^{n=1}} < \frac{\varepsilon}{2}. \tag{1.64}$$

Now, on the left-hand side of [1.64], we have $\bigcup_{n=1}^{p_0} G_n = U_{p_0}$ and we can replace $\lambda_{\text{op}}\left(U_{p_0} \setminus \bigcup_{n=1}^{p_0} K_n\right)$ with $\lambda_{\text{op}}\left(U_{p_0}\right) - \lambda_{\text{cp}}\left(\bigcup_{n=1}^{p_0} K_n\right)$. So [1.64] amounts to $\lambda_{\text{op}}\left(U_{p_0}\right) - \lambda_{\text{cp}}\left(\bigcup_{n=1}^{p_0} K_n\right)$, and we ahve thus obtained the inequality

$$\lambda_{\text{op}}\left(U_{p_0}\right) < \lambda_{\text{cp}}\left(\bigcup_{n=1}^{p_0} K_n\right) + \frac{\varepsilon}{2}. \tag{1.65}$$

Finally, by putting together [1.65] with the inequality from Claim 2, we find that

$$\lambda_{\text{op}}(U) < \lambda_{\text{op}}\left(U_{p_0}\right) + \frac{\varepsilon}{2} < \lambda_{\text{cp}}\left(\bigcup_{n=1}^{p_0} K_n\right) + \varepsilon,$$

which is what the condition (c) was asking for.

QED

Proposition 1.35.

$A \in \mathcal{M}_{\text{bdd}}$ with

$$\lambda_{\text{bdd}}(A) = \sum_{n=1}^{\infty} \lambda_{\text{bdd}}(A_n).$$

Proof. To show that $A \in \mathcal{M}_{\text{bdd}}$, we have to prove that for every $\varepsilon > 0$, there are a compact set K and a bounded open set G such that $K \subseteq A \subseteq G$ and that $\lambda_{\text{op}}(G) < \lambda_{\text{cp}}(K) + \varepsilon$. This is found by using Lemma 1.34: in the notation of the lemma, the needed open set is U and the needed compact set is $\bigcup_{n=1}^{p_0} K_n$. This verifies $A \in \mathcal{M}_{\text{bdd}}$.

Now, denote $l = \sum_{n=1}^{\infty} \lambda_{\text{bdd}}(A_n)$. We then have

$$l = \lim_{p\to\infty} \sum_{n=1}^{p} \lambda_{\text{bdd}}(A_n) = \sup_{p\in\mathbb{N}} \sum_{n=1}^{p} \lambda_{\text{bdd}}(A_n),$$

hence the inequality $\lambda_{\text{bdd}}(A) \geq l$ will follow if we can prove that

$$\forall p \in \mathbb{N}\left[\lambda_{\text{bdd}}(A) \geq \sum_{n=1}^{p} \lambda_{\text{bdd}}(A_n)\right]. \tag{1.66}$$

But the inequality [1.66] follows easily from known properties of λ_{bdd}:

$$\begin{aligned}\sum_{n=1}^{p} \lambda_{\text{bdd}}(A_n) &= \lambda_{\text{bdd}}\left(\bigcup_{n=1}^{p} A_n\right) && \text{since } \lambda_{\text{bdd}} \text{ is additive}\\ &\le \lambda_{\text{bdd}}(A) && \text{since } \lambda_{\text{bdd}} \text{ is increasing}\end{aligned}$$

for all $p \in \mathbb{N}$. This concludes that $\lambda_{\text{bdd}}(A) \ge l$.

It remains to show that $\lambda_{\text{bdd}}(A) \le l$. For this inequality, it is convenient to resort to the trick of *leaving an ε of a room*: it is sufficient to prove that

$$\forall \varepsilon > 0 \left[\lambda_{\text{bdd}}(A) < l + \varepsilon\right]. \tag{1.67}$$

We then fix $\varepsilon > 0$, for which we verify $\lambda_{\text{bdd}}(A) < l + \varepsilon$. In connection to this ε, we appeal to Lemma 1.34, and we take in the compact sets $K_1, \ldots, K_{p_0}$ and the bounded open set U provided by the lemma.

$$\begin{aligned}\lambda_{\text{bdd}}(A) &\le \lambda_{\text{bdd}}(U) && \text{since } \lambda_{\text{bdd}} \text{ is increasing}\\ &= \lambda_{\text{op}}(U)\\ &< \varepsilon + \lambda_{\text{cp}}\left(K_1 \cup \cdots \cup K_{p_0}\right) && \text{by definition of } K_1, \ldots, K_{p_0}\\ &= \varepsilon + \lambda_{\text{bdd}}\left(K_1 \cup \cdots \cup K_{p_0}\right)\\ &= \varepsilon + \sum_{n=1}^{p_0} \lambda_{\text{bdd}}(K_n) && \text{since } \lambda_{\text{bdd}} \text{ is additive}\\ &\le \varepsilon + \sum_{n=1}^{p_0} \lambda_{\text{bdd}}(A_n) && \text{since } \lambda_{\text{bdd}} \text{ is increasing}\\ &= \varepsilon + l.\end{aligned}$$

We have now verified both inequalities $\lambda_{\text{bdd}}(A) \ge l$ and $\lambda_{\text{bdd}}(A) \le l$, and this concludes the proof.

QED

We now return to our goal of proving that $\mathcal{M}$ is a σ-algebra of subsets of $\mathbb{R}$.

Lemma 1.36.
Let $\{M_n\}_{n=1}^{\infty}$ be a collection of pairwise disjoint sets from $\mathcal{M}$. Then $\bigcup_{n=1}^{\infty} M_n \in \mathcal{M}$.

Proof. Denote $M = \bigcup_{n=1}^{\infty} M_n$. In order to prove that $M \in \mathcal{M}$, it suffices to check that (by Proposition 1.29)

$$\forall k \in \mathbb{N} \left[M \cap (-k, k) \in \mathcal{M}_{\text{bdd}}\right]. \tag{1.68}$$

We thus fix $k \in \mathbb{N}$, for which we verify that $M \cap (-k, k) \in \mathcal{M}_{\text{bdd}}$.

Define $A_n = M_n \cap (-k, k)$ for all $n \in \mathbb{N}$. Every A_n belongs to $\mathcal{M}_{\text{bdd}}$ and is obviously contained in $(-k, k)$. Moreover, for distinct indices $m, n \in \mathbb{N}$ we have that $A_m \cap A_n = \varnothing$, due to the fact that $A_m \subseteq M_m, A_n \subseteq A_n$, and we have as hypothesis that $M_m \cap M_n = \varnothing$. It follows that Proposition 1.35 can be applied to the collection $\{A_n\}_{n=1}^{\infty}$: it gives us the conclusion that $\bigcup_{n=1}^{\infty} A_n \in \mathcal{M}_{\text{bdd}}$. But

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} \left(M_n \cap (-k, k)\right) = (-k, k) \cap \bigcup_{n=1}^{\infty} M_n = M \cap (-k, k).$$

So we have obtained that $M \cap (-k, k) \in \mathcal{M}_{\text{bdd}}$, as required.

QED

Proposition 1.37.
Let $\{M_n\}_{n=1}^{\infty}$ be a collection of sets in $\mathcal{M}$. Then $\bigcup_{n=1}^{\infty} M_n \in \mathcal{M}$.

Proof. Unlike what we had in Lemma 1.36, here there is no assumption that the sets M_n would be pairwise disjoint. But we can nevertheless fall back on Lemma 1.36 by using the following trick: consider the collection $\{M_n'\}_{n=1}^{\infty}$ defined by

$$M_n' = M_n \setminus \bigcup_{m=1}^{n-1} M_m$$

for all $n \in \mathbb{N}$. The fact that $\mathcal{M}$ is an algebra of sets ensures that $M'_n \in \mathcal{M}$ for all $n \in \mathbb{N}$. On the other hand, basic verifications involving set operations show us that $\bigcup_{n=1}^{\infty} M'_n = \bigcup_{n=1}^{\infty} M_n$ and that $M'_n \cap M'_m = \varnothing$ for all distinct $m, n \in \mathbb{N}$. Thus Lemma 1.36 applies to the collection $\{M'_n\}_{n=1}^{\infty}$, and concludes that $\bigcup_{n=1}^{\infty} M_n = \bigcup_{n=1}^{\infty} M'_n \in \mathcal{M}$.

QED

Corollary 1.37.1.
$\mathcal{M}$ is a σ-algebra of subsets of $\mathbb{R}$.

See Proposition 1.31, 1.37

Lastly, we are going to show that λ is a positive measure.

Lemma 1.38.
Let $\{M_n\}_{n=1}^{\infty}$ be a collection of pairwise disjoint sets from $\mathcal{M}$ and let $l = \sum_{n=1}^{\infty} \lambda(M_n) \in [0, \infty]$. On the other hand, consider the set $M = \bigcup_{n=1}^{\infty} M_n$, which is still in $\mathcal{M}$. We have the inequality

$$l \leq \lambda(M). \tag{1.69}$$

Proof. Let us write explicitly that

$$l = \lim_{p\to\infty} \sum_{n=1}^{p} \lambda(M_n) = \sup_{p\in\mathbb{N}} \sum_{n=1}^{p} \lambda(M_n).$$

Since the supremum is a *least* upper bound, we see that [1.67] will follow as soon as we prove that

$$\forall p \in \mathbb{N} \left[\sum_{n=1}^{p} \lambda(M_n) \leq \lambda(M) \right]. \tag{1.70}$$

And indeed, for each $p \in \mathbb{N}$, the inequality in [1.68] follows from Proposition 1.33.

QED

Lemma 1.39.
Consider the setting from Lemma 1.38. We also have:

$$l \geq \lambda(M). \tag{1.71}$$

Proof. Since $\lambda(M) = \lim_{k\to\infty} \lambda_{\text{bdd}}(M \cap (-k, k))$, it suffices to prove that

$$\forall k \in \mathbb{N} \left[\lambda_{\text{bdd}}(M \cap (-k, k)) \leq l \right]. \tag{1.72}$$

We fix $k \in \mathbb{N}$ for which we verify that $\lambda_{\text{bdd}}(M \cap (-k, k)) \leq l$ holds.

Let us resort again to the tick of the ε *of room*. For the k that we fixed, the desired inequality will follow if we can show that

$$\forall \varepsilon > 0 \left[\lambda_{\text{bdd}}(M \cap (-k, k)) - \varepsilon < l \right]. \tag{1.73}$$

Hence, in addition to the fixed k, let us also fix $\varepsilon > 0$, and go for the proof of [1.73].

Observe that $M \cap (-k, k) = (-k, k) \cap \bigcup_{n=1}^{\infty} M_n = \bigcup_{n=1}^{\infty} (M_n \cap (-k, k))$, where the sets $M_n \cap (-k, k)$ are in $\mathcal{M}_{\text{bdd}}$, pairwise disjoint, and contained in $(-k, k)$. Hence we are precisely in the situation where Proposition 1.35 can be applied. The said proposition tells us that

$$\lambda_{\text{bdd}}(M \cap (-k, k)) = \sum_{n=1}^{\infty} \lambda_{\text{bdd}}(M_n \cap (-k, k)) = \lim_{p\to\infty} \sum_{n=1}^{p} \lambda_{\text{bdd}}(M_n \cap (-k, k)). \tag{1.74}$$

In [1.74] we are dealing with a finite limit of an increasing sequence in $[0, \infty)$, so there is no problem to find a $p_0 \in \mathbb{N}$ such that

$$\sum_{n=1}^{p_0} \lambda_{\text{bdd}}(M_n \cap (-k, k)) > \lambda(M \cap (-k, k)) - \varepsilon.$$

Using this p_0 we write that

$$l = \sum_{n=1}^{\infty} \lambda\left(M_n\right) \geq \sum_{n=1}^{p_0} \geq \sum_{n=1}^{p_0} \lambda_{\text{bdd}}\left(M_n \cap (-k, k)\right) > \lambda\left(M \cap (-k, k)\right) - \varepsilon,$$

and this leads to the inequality [1.73] we had been left to prove.

QED

Corollary 1.39.1.

The map $\lambda : \mathcal{M} \to [0, \infty]$ is σ-additive, and thus a positive measure.

See Corollary 1.37.1, Lemma 1.38, 1.39

II. Measurable Functions

1. Non-measurable Sets

As a result of the effort made in the preceding section, we now have in hand

(a) a collection $\mathcal{M}$ of *Lebesgue measurable* subsets of $\mathbb{R}$; and

(b) a length-measuring map $\lambda : \mathcal{M} \to [0, \infty]$, called *Lebesgue measure*.

Moreover, we checked that $\mathcal{M}$ is a σ-algebra and that λ is a positive measure. In connection to these general notions, here is a bit of additional terminology which is part of the measure and integration theory.

Def'n 2.1. **Measurable Space**, **Measure Space**
Let X be a nonempty set.

(a) When $\mathcal{A}$ is a σ-algebra on X, we call the pair $(X, \mathcal{A})$ a ***measurable space***.

(b) If $\mu : \mathcal{A} \to [0, \infty]$ is a positive measure in addition, then we say $(X, \mathcal{A}, \mu)$ is a ***measure space***.

In the terms set by Def'n 2.1, we have that $(\mathbb{R}, \mathcal{M})$ is a measurable space and that $(\mathbb{R}, \mathcal{M}, \lambda)$ is a measure space.

The notion of measurable space offers a convenient framework for talking about *measurable functions*, which is pretty much the topic of this section. Then the notion of measure space offers a convenient framework for discussing about *integrable functions*, which is what we will do when we are done with setting the basics for measurable functions.

Knowing that we are dealing with a σ-algebra is, without a doubt, a rather comfortable thing for what we do. Indeed, since all the open sets are min $\mathcal{M}$, we know that we can start from open sets and we can do any kind of set operations we want, involving *countably many* sets at a time, and the set M resulting out of these operations will still be in $\mathcal{M}$, so we can fearlessly talk about the length (i.e. Lebesgue measure) $\lambda(M)$. Thus $\lambda(M)$ can surely be considered when M is a closed set, then also for sets that are of type G_δ, of type F_σ, and so on.

In connection to the above, here are the two questions that we consider in this and next subsections.

- *Question 1. Couldn't it actually be that **every** subset of $\mathbb{R}$ is Lebesgue measurable?*

The answer will be negative. But, exactly because the σ-algebra $\mathcal{M}$ is so comfortable to deal with, it's not that easy to produce an example of non-measurable set.

- *Question 2. Is $\mathcal{M}$ the **minimal** σ-algebra of subsets of $\mathbb{R}$ that we get by starting from open sets?*

Once again, the answer is negative. There is a slightly smaller σ-algebra, called the *Borel σ-algebra*, which has this minimality property. This is discussed in the next subsection.

The next lemma shows the blueprint we will use in order to produce an example of non-measurable set.

Def'n 2.2. **Translation** of a Set
Let $S \subseteq \mathbb{R}, x \in \mathbb{R}$. We define the ***translation*** of S by x, denoted as $S + x$, by

$$S + x = \{s + x : s \in S\}.$$

Exercise 2.1.
Let $M \in \mathcal{M}$. Then for all $x \in \mathbb{R}$, show that $M + x \in \mathcal{M}$ and that $\lambda(M + x) = \lambda(M)$.

Assignment

Lemma 2.1. A Criterion for Non-measurability
Let $S \subseteq (0,1)$. If there exists a countable set $\{q_n\}_{n=1}^{\infty}$ of numbers in $(-1,1)$ such that

$$\forall m, n \in \mathbb{N}\,[m \neq n \implies (S + q_m) \cap (S + q_n) = \varnothing], \tag{2.1}$$

and that

$$\bigcup_{n=1}^{\infty}(S+q_n)\supseteq(0,1)\,, \tag{2.2}$$

then $S\notin\mathcal{M}$.

Proof. Let us assume, for contradiction, that $S\notin\mathcal{M}$, and let us denote $\lambda(S)=c\in[0,\infty]$. Since it is given that, we can actually be sure that

$$c=\lambda(S)\le\lambda((0,1))=1,$$

hence that $c\in[0,1]$.

For every $n\in\mathbb{N}$, denote $M_n=S+q_n$. Sicne we assumed that $S\in\mathcal{M}$, it follows that every M_n is in $\mathcal{M}$, with $\lambda(M_n)=\lambda(S)=c$, in view of Exercise 2.1. Another observation about the sets $M-n$ is that we have

$$\forall n\in\mathbb{N}\,[M_n\subseteq(-1,2)]\,. \tag{2.3}$$

This is because $S\subseteq(0,1)$ and each q_n is in $(-1,1)$.

Now consider the set $M=\bigcup_{n=1}^{\infty}M_n$. Then $m\in\mathcal{M}$, because the M_n's are from $\mathcal{M}$, and $\mathcal{M}$ is a σ-algebra. From [2.3] it follows that $M\subseteq(-1,2)$ and, on the other hand, the hypothesis [2.2] of the lemam says precisely that $M\supseteq(0,1)$. This yields some estimates on the Lebesgue measure of M: sicne $(0,1)\subseteq M\subseteq(-1,2)$, we have $\lambda((0,1))\le\lambda(M)\le\lambda((-1,2))$. We thus find that

$$1\le\lambda(M)\le3. \tag{2.4}$$

We are very close to a contradiction. Indeed, recall that $M=\bigcup_{n=1}^{\infty}M_n$, where the hypothesis [2.1] tells us that $M_m\cap M_n=\varnothing$ for all distinct $n,m\in\mathbb{N}$. So then by the σ-additivity of the Lebesgue measure implies that

$$\lambda(M)=\sum_{n=1}^{\infty}\lambda(M_n)=\lim_{p\to\infty}\sum_{n=1}^{p}\lambda(M_n)=\lim_{p\to\infty}pc=\begin{cases}0 & \text{if } c=0\\ \infty & \text{if } 0<c\le1\end{cases}.$$

Either way, this contradicts with the estimate [2.4].

Thus we conclude that S is not Lebesgue measurable.

QED

Now, of course, we have to ask: *how do we find a set S and a set of numbers $\{q_n\}_{n=1}^{\infty}$ which satisfy the hypothesis of Lemma 2.1?* It turns out that we can arrange that to happen in connection to an equivalence relation, which we define on $(0,1)$, as follows. For all $x,y\in(0,1)$, write

$$x\sim y\iff x-y\in\mathbb{Q}\,.$$

Verifying $\sim$ is an equivalence relation amounts to the fact that $(\mathbb{Q},+)$ is an additive group.

Now, having an equivalence relation $\sim$ *splits* $(0,1)$ into a disjoint union of equivalence classes with respect to $\sim$:

$$(0,1)=\bigcup_{i\in I}E_i$$

with $E_i\cap E_j=\varnothing$ for all distinct $i,j\in I$.

From every equivalence class E_i, we *choose* a point, say $s_i\in E_i$, by the axiom of choice. Let us define

$$S=\{s_i:i\in I\}\,,$$

which indeed is a subset of $(0,1)$.

Proposition 2.2.
Consider S defiend as above and let $(q_n)_{n=1}^{\infty}$ be an enumeration of the rational numbers in $(-1,1)$. Then S and $\{q_n\}_{n=1}^{\infty}$ satisfy the hypotheses of Lemma 2.1.

Proof. We must check that the conditions [2.1] and [2.2] from Lemma 2.1 are satisfied.

For [2.1], we pick some distinct indices $n, m \in \mathbb{N}$, for which we verify that $(S + q_m) \cap (S + q_n) = \varnothing$.

Assume for contradiction that there is $x \in (S + q_m) \cap (S + q_n)$. Then there are $s', s'' \in S$ such that $x = s' + q_m = s'' + q_n$. In view of the definition of S, these numbers s' and s'' have to be of the form s_i and s_j for some $i, j \in I$, respectively. We thus have equalities

$$x = s_i + q_m = s_j + q_n. \tag{2.5}$$

But from [2.5], combined how the elements of S were selected to represent the equivalence classes of $\sim$, it follows that

$$s_i - s_j = q_n - q_m \in \mathbb{Q} \implies s_i \sim s_j \implies i = j.$$

So then $s_i = s_j$, and from the equality $s_i + q_m = s_j + q_n$ we infur that $q_m = q_n$, which is a contradiction.

We thus conclude that $(S + q_m) \cap (S + q_n) = \varnothing$.

For [2.2], inview of what $\bigcup_{n=1}^{\infty} (S + q_n)$ means, the verification that needs to be done here is this: for all $x \in (0,1)$, there exists $i \in I$ such that $x \in E_i$. In the equivalence class E_i we have selected representative s_i; so then, it is the case that $x - s_i \in \mathbb{Q}$. We also observe that, since $x, s_i \in (0,1)$, we have $-1 < x - s_i < 1$. Thus $x - s_i \in \mathbb{Q} \cap (-1,1)$, and there there exists an $n \in \mathbb{N}$ such that $x - s_i = q_N$. This n is exactly what we need, since $x - s_i = q_n$ implies $x = s_i + q_n \in S + q_n$.

QED

As a consequence to Proposition 2.2, S is not Lebesgue measurable.

2. Borel σ-algebras

In this subsection, we discuss a more general method to produce σ-algebras: start with an arbitrary collection $\mathcal{C}$ of subsets of a set X, and consider *the smallest possible σ-algebra which contains $\mathcal{C}$*. We then use this to construct the said Borel σ-algebra.

The precise description of what the above means appears in Proposition 2.4. Before stating that proposition, we record, in the next lemma, a simple observation coming out directly from the definition of a σ-algebra.

Lemma 2.3.

Let X be a nonempty set and let $\{\mathcal{A}_i\}_{i \in I}$ be a family of σ-algebras of subsets of X. Denote $\mathcal{A} = \bigcap_{i \in I} \mathcal{A}_i$. Then $\mathcal{A}$ is a σ-algebra.

Proof. We verify three things.

(a) Since $\varnothing, X \in \mathcal{A}_i$ for all $i \in I$, $\varnothing, X \in \mathcal{A}$.

(b) Let $N, M \in \mathcal{A}$. Then for all $i \in I$, $N, M \in \mathcal{A}_i$, so that $N \setminus M \in \mathcal{A}_i$.

(c) Let $\mathcal{C}$ be a countable collection of sets in $\mathcal{A}$. Then $\mathcal{C} \subseteq \mathcal{A}_i$ for all $i \in I$, so that $\bigcup \mathcal{C} \in \mathcal{A}_i$. Thus $\bigcup \mathcal{C} \in \mathcal{A}$.

QED

Proposition 2.4.

Let X be a nonempty set and let $\mathcal{C}$ be a collction of subsets of X. Then there exists a unique collection $\mathcal{A}_0$ of subsets of X such that

(a) $\mathcal{A}_0$ is a σ-algebra of subsets of X containing $\mathcal{C}$; and

(b) every σ-algebra containing $\mathcal{C}$ contains $\mathcal{A}_0$ also.

Proof. We first show that such $\mathcal{A}_0$ exists.

Let $\{\mathcal{A}_i\}_{i \in I}$ be the family of all the σ-algebras of subsets of X which contain the given $\mathcal{C}$.[1] We put

$$\mathcal{A}_0 = \bigcap_{i \in I} \mathcal{A}_i,$$

which is a σ-algebra of subset of X by Lemma 2.3 and contains $\mathcal{C}$ since every $\mathcal{A}_i$ contains $\mathcal{C}$. Hence $\mathcal{A}_0$ satisfies (a).

On the other hand, let $\mathcal{A}$ be a σ-algebra of subsets of X containing $\mathcal{C}$. Then $\mathcal{A} = \mathcal{A}_j$ for some $j \in I$, so that

$$\mathcal{A} = \mathcal{A}_j \supseteq \bigcap_{i \in I} \mathcal{A}_i = \mathcal{A}_0 .$$

Thus $\mathcal{A}_0$ satisfies (b) as well.

For the uniqueness, let $\mathcal{A}_0$ be defined as above. Let $\mathcal{A}_0'$ be a σ-algebra of subsets of X which also satisfies (a), (b). We have to prove that $\mathcal{A}_0' = \mathcal{A}_0$.

But (b) applies to both $\mathcal{A}_0, \mathcal{A}_0'$, so that $\mathcal{A}_0 \subseteq \mathcal{A}_0'$ and $\mathcal{A}_0' \subseteq \mathcal{A}_0$. Thus we conclude $\mathcal{A}_0' = \mathcal{A}_0$, as required.

[1] We note that 2^X, the power set of X, is a σ-algebra of subsets of X. Hence such family is nonempty.

QED

Def'n 2.3. σ-algebra **Generated** by a Collection

Let X be a nonempty set and let $\mathcal{C}$ be a collection of subsets of X. Then the σ-algebra $\mathcal{A}_0$ found in Proposition 2.4 is called the σ-algebra of subsets of X ***generated*** by $\mathcal{C}$, which we shall denote as $\sigma\text{-Alg}(\mathcal{C})$.

The argument used to prove Proposition 2.4 is qutie far from being constructive. Nevertheless, we will see that useful things can be proved about $\sigma\text{-Alg}(\mathcal{C})$, by just exploiting the conditions in Proposition 2.4. An easy example of how this goes is provided by the next exercise.

Exercise 2.2.

Let X be a nonempty set and let $\mathcal{C}_1$ and $\mathcal{C}_2$ be collections of subsets of X such that $\mathcal{C}_1 \subseteq \mathcal{C}_2$. Prove that $\sigma\text{-Alg}(\mathcal{C}_1) \subseteq \sigma\text{-Alg}(\mathcal{C}_2)$.

Proof. It suffices to note that $\sigma\text{-Alg}(\mathcal{C}_2) \supseteq \mathcal{C}_2 \supseteq \mathcal{C}_1$, so $\sigma\text{-Alg}(\mathcal{C}_2)$ contains $\sigma\text{-Alg}(\mathcal{C}_1)$, the smallest σ-algebra containing $\mathcal{C}_1$.

QED

Def'n 2.4. **Borel σ-algebra**

We call $\sigma\text{-Alg}(\mathcal{T})$, the σ-algebra generated by the open subsets of $\mathbb{R}$, the ***Borel σ-algebra*** of $\mathbb{R}$, denoted as $\mathcal{B}$.

$\mathcal{B}$ is the smallest σ-algebra of subsets of $\mathbb{R}$ which contains all the open sets. In order to prove things about Borel sets, we will typically just fall back on the description of $\mathcal{B}$ via the properties (a), (b) from Proposition 2.4.

In connection to the above, observe that it is possible (and sometimes convenient) to approach the Borel σ-algebra $\mathcal{B}$ by using a collection of generators different from $\mathcal{T}$. Here is an instructive exercise on these lines, which says that *one can generate $\mathcal{B}$ by using compact sets*.

Exercise 2.3.

Recall that we denote $\mathcal{K}$ to be the collection of compact subsets of $\mathbb{R}$. Prove that $\sigma\text{-Alg}(\mathcal{K}) = \mathcal{B}$.

tl;dr

One thing which we clearly have at this point is that $\mathcal{B} \subseteq \mathcal{M}$. This is because $\mathcal{M}$ is a σ-algebra which contains $\mathcal{T}$, while $\mathcal{B}$ is the *minimal* σ-algebra which contains $\mathcal{T}$. In order to determine that $\mathcal{M} \neq \mathcal{B}$, we need to work a bit more.

For the time being, we just record the fact that, since $\mathcal{B} \subseteq \mathcal{M}$, we can in any case restrict the Lebesgue measure λ to $\mathcal{B}$, which will clearly give us a positive measure on $\mathcal{B}$. We are thus getting a measure space $(\mathcal{R}, \mathcal{B}, \lambda)$, which will turn out to be of great interest for us.

When we look at how $\mathcal{M}$ and λ appeared in our considerations, we see that they were *packed together*. For instance, look at where $\mathcal{M}_{\text{bdd}}$ was introduced. We see that we were using there λ_{op} and λ_{cp}, the preliminary instances of λ we had developed for open and for compact sets.

So the definition of $\mathcal{M}$ is intimately related to the notion of length. The definition of $\mathcal{B}$ is not like that. In order to define $\mathcal{B}$ we only need to know what are the open subsets of $\mathbb{R}$. This immediately prompts the thought that we can define a Borel σ-algebra associated to any topological space. Later in the course it will be useful to play with Borel subsets of such metric spaces X, for instance make X be the unit circle in the complex plane.

Def'n 2.5. **Borel σ-algebra** of a Topological Space

Let $(X, \mathcal{T})$ be a topological space. We call $\sigma\text{-Alg}(\mathcal{T})$ the ***Borel σ-algebra*** on X, which is denoted as $\mathcal{B}_X$.[1] The elements of $\mathcal{B}_X$ are called the ***Borel sets***.

[1] We should really write $\mathcal{B}_{(X,\mathcal{T})}$ instead, since the definition heavily depends on $\mathcal{T}$. However, often times $\mathcal{T}$ is well-understood, so we shall write $\mathcal{B}_X$ for convenience.

As before, we are writing $\mathcal{B}$ to mean $\mathcal{B}_{\mathbb{R}}$.

Suppose that our metric space is $[0,1]$, endowed with the usual distance $d(x,y) = |x-y|$ for all $x, y \in [0,1]$. This is a compact metric space. We can talk about the collection $\mathcal{T}_{[0,1]}$ of subsets of $[0,1]$ relatively open in $[0,1]$, and we can then consider the corresponding Borel σ-algebra:

$$\mathcal{B}_{[0,1]} = \sigma\text{-Alg}\left(\mathcal{T}_{[0,1]}\right).$$

Here is a natural question that pops up in connection to this. Namely, since $[0,1] \subseteq \mathbb{R}$ and we already have $\mathcal{B}_{\mathbb{R}}$, *why don't we actually work with the collection* $\hat{\mathcal{B}}_{[0,1]}$ *of subsets of* $[0,1]$ *defiend by*

$$\hat{\mathcal{B}}_{[0,1]} = \{M \in \mathcal{B}_{\mathbb{R}} : M \subseteq [0,1]\}. \tag{2.6}$$

The good news is that $\hat{\mathcal{B}}_{[0,1]}$ coincides with $\mathcal{B}_{[0,1]}$ (which is a special case of the situation considered in Proposition 2.6). This means [2.6] can be used as a description of the Borel σ-algebra of $[0,1]$.

Here is an example which lives in the complex plane, and will be important in the final part of this course: consider the unit circle

$$\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$$

on the complex plane, endowed with the usual distance between complex numbers, $d(z,w) = |z-w|$ for all $z, w \in \mathbb{T}$. Then $(\mathbb{T}, d)$ is a compact metric space, which has its collection $\mathcal{T}_{\mathbb{T}}$ of open sets, and the corresponding Borel σ-algebra $\mathcal{B}_{\mathbb{T}} = \sigma\text{-Alg}(\mathcal{T}_{\mathbb{T}})$.

Proposition 2.5.

Let $(X, \mathcal{A})$ be a measurable space. Let X_0 be a nonempty set in $\mathcal{A}$ and let

$$\mathcal{A}_0 = \{A \in \mathcal{A} : A \subseteq X_0\}.$$

Then $\mathcal{A}_0$ is a σ-algebra of subsets of X_0.

Quite Clear!

Def'n 2.6. **Restriction** of a σ-algebra

Consider the setting of Proposition 2.5. We call $(X_0, \mathcal{A}_0)$ the ***restriction*** of $(X, \mathcal{A})$ to X_0.

Proposition 2.6.

Let (X, d) be a metric space. Consider the collection of open sets $\mathcal{T}_X$ of (X, d) and the Borel σ-algebra $\mathcal{B}_X$.

Let $X_0 \in \mathcal{B}_X$ be nonempty and let $(X_0, \mathcal{A}_0)$ be the restriction of the measurable space $(X, \mathcal{B}_X)$ to X_0. On the other hand, let $d_0 = d|_{X_0}$, the restriction of the metric on X, which is a metric on X_0. Consider the Borel σ-algebra $\mathcal{B}_{X_0}$.

Then $\mathcal{B}_{x_0} = \mathcal{A}_0$.

tl;dr

We now return to the problem showing $\mathcal{B} \neq \mathcal{M}$. That is, there exists $M \subseteq \mathbb{R}$ that is Lebesgue measurable but is not a Borel set. The idea is that we will find such M to be *negligible*. Recall the definition of such sets.

Recall 2.7. **Negligible** Set

We say $N \subseteq \mathbb{R}$ is ***negligible*** if N is Lebesgue measurable with measure 0.

From assignments, we have an alternative description of negligible sets, which is recorded in the next proposition.

Proposition 2.7.

Let $N \subseteq \mathbb{R}$. The following are equivalent.

(a) N is negligible.

(b) For all $\varepsilon > 0$, there exists open $G \subseteq \mathbb{R}$ that contains N with $\lambda_{\text{op}}(G) < \varepsilon$.

Assignment!

A benefit of the characterization appearing in Proposition 2.7 is that it has the following immediate consequence.

Corollary 2.7.1.
Every subset of a negligible set is negligible.

Proof by Inspection!

Let us put to use the work on the tenary Canor set on assignments.

Corollary 2.7.2.
Let C be the tenary Cantor set. Then every subset of C is measurable.

Proof. We saw that C is negligible, which means every subset of C is negligible (and in particular, measurable) by Corollary 2.7.1.

QED

Exercise 2.4.
Let $\mathcal{N}$ be a countable collection of negligible subsets of $\mathbb{R}$. Prove that $\bigcup \mathcal{N}$ is also negligible.

Proof. We may assume $\mathcal{N}$ is countably infinite, so denote $\mathcal{N} = \{N_n\}_{n=1}^{\infty}$. Suppose $\varepsilon > 0$, for which we verify that there is open $G \subseteq \mathbb{R}$ containing $\bigcup \mathcal{N}$ with $\lambda_{\text{op}}(G) < \varepsilon$.

For all $n \in \mathbb{N}$, let G_n be an open set containing N_n with $\lambda_{\text{op}}(G_n) < \frac{\varepsilon}{2^{n+1}}$. Such G_n exists by Proposition 2.7 since N_n is negligible. Now define

$$G = \bigcup_{n=1}^{\infty} G_n,$$

which is open as a union of open sets. Moreover, G contains $N = \bigcup_{n=1}^{\infty} N_n$ since each G_n contains N_n, with

$$\lambda(G) \le \sum_{n=1}^{\infty} \lambda(G_n) \le \sum_{n=1}^{\infty} \frac{\varepsilon}{2^{n+1}} = \frac{\varepsilon}{2} \sum_{n=1}^{\infty} \frac{1}{2^n} = \frac{\varepsilon}{2} < \varepsilon,$$

as required.

QED

Our goal of proving the existence of measurable non-Borel sets will then achieved with the following proposition.

Proposition 2.8.
Let C be the tenary Cantor set. Then there is a subset of C which is not Borel.

Proof Postponed

The proof of Proposition 2.8 goes a bit outside the tools that we have available at this moment. We list below three facts that we combine in order to get this proof.

Fact 2.9.

(a) Let $K \subseteq [0,1]$ be a compact non-negligible set. Then there is a non-measurable subset of K.

(b) Let K_1, K_2 be nonempty compact subsets of $\mathbb{R}$ and let $\varphi : K_1 \to K_2$ be a homeomorphism. Consider the Borel σ-algebras $\mathcal{B}_{K_1}$ and $\mathcal{B}_{K_2}$ associated to K_1 and K_2, respectively. Then $\varphi(\mathcal{B}_{K_1}) \subseteq \mathcal{B}_{K_2}$.

(c) Let C be the tenary Cantor set. Then there exists a continuous injection $\varphi : C \to [0,1]$ such that the compact set $K = \varphi(C)$ has a nonzero measure.

Here are some comments concerning these facts.

(a) is an (not-so-immediate) upgrade of the argument shown when we constructed a non-measurable set.

(b) and (c) refer to functions between measurable spaces, and are best treated after we discuss a bit about such functions. In fact the function φ of (c) will be very interesting to look at. It must *stretch* distances between points quite widely if it succeeds to start with the negligible compact set C and map it onto a non-negligible set $\varphi(C)$.

We now demonstrate how Proposition 2.8 can be proved if we assume Fact 2.9.

Proof of Proposition 2.8

Let us assume that every subset of C is a Borel set for contradiction. Then the restriction of the measurable space $(\mathbb{R}, \mathcal{B})$ to C is $\left(C, 2^C\right)$. Thus when we view C as a compact metric space, its Borel σ-algebra $\mathcal{B}_C$ comes out as $\mathcal{B}_C = 2^C$.

Now consider the homeomorphism $\varphi : C \to K$ that is provided to us by (c) of Fact 2.9.[1] We recall that $\lambda(K) > 0$.

- *Claim 1.* $\mathcal{B}_K = 2^K$.

  Proof. What we have to prove is that every subset of T is in $\mathcal{B}_K$ (note that $\mathcal{B}_k \subseteq 2^K$ is immediate). To that end, fix $T \subseteq K$.

  Let $S = \varphi^{-1}(T) \subseteq C$. Since $\mathcal{B}_C = 2^C$, we have that $S \in \mathcal{B}_C$. Then (b) of Fact 2.9 implies $\varphi(S) \in \mathcal{B}_K$. But $\varphi(S) = T$, so that $T \in \mathcal{B}_K$.

But by using Proposition 2.6, Claim 1 can be read as saying that *every subset of K belongs to the Borel σ-algebra $\mathcal{B}$.* Since K is a compact subset of $[0,1]$ with a nonzero measure, we reached a contradiction with (a) of Fact 2.9.

[1]Note that (c) of Fact 2.9 gives only a continuous *injection* $\varphi : C \to [0,1]$ such that $\varphi(C)$ has a nonzero measure. By restricting its codomain to $\varphi(C)$, we obtain a homeomorphism, so that we can apply (b) of Fact 2.9.

QED

If one looks into the literature, one will also find another approach to proving Proposition 2.8, which is entirely set-theoretic: one can compare the infinite cardinalities of $\mathcal{B}_C$ and 2^C, and arrive to the conclusion that $|\mathcal{B}_C| < \left|2^C\right|$. Hence in particular $\mathcal{B}_C \subset 2^C$, as Proposition 2.8 is stating.

Finnaly, here is a little proposition driving the idea that, while $\mathcal{B}$ is a proper subcollection of $\mathcal{M}$, the difference between two really is about how they handle negligible sets. From the point of view of integration theory, this is not such a crucial difference, since negligible sets can most of the time be ignored in the process of integration.

Proposition 2.10.

Let M be a measurable set. Then there is $B \subseteq M$ such that B is a Borel set and $N = M \setminus B$ is negligible.

Proof. When M is bounded, from assignment we know that $M = F \cup N$ for some F_σ set F and some negligible N. Since every F_σ set is Borel (as a countable union of closed sets), we reach the desired conclusion by taking $B = F$.

Suppose M is not bounded. For all $n \in \mathbb{N}$, let $M_n = M \cap (-n, n)$, which is a bounded measurable set. Then the observation in the preceding paragraph can be applied to every M_n; it gives a set $B_n \subseteq M_n$ such that $B_n \in \mathcal{B}$ and such that $N_n = M_n \setminus B_n$ is negligible. We then take

$$B = \bigcup_{n=1}^{\infty} B_n$$

and

$$N = M \setminus B.$$

Then a direct inspection gives that $B \subseteq \bigcup_{n=1}^{\infty} M_n = M$ and that $N \subseteq \bigcup_{n=1}^{\infty} \left(M_n \setminus B_n\right) = \bigcup_{n=1}^{\infty} N_n$. Hence $B \in \mathcal{B}$ since $\mathcal{B}$ is closed under countable intersections and N is negligible as a countable union of negligible sets (Exercise 2.4). Thus B and N achieve a decomposition of M as required by the proposition.

QED

3. Measurable Functions

For this subsection, we will momentarily forget about λ, and just look at the measurable space $(\mathbb{R}, \mathcal{B})$. We are interested in functions $f : \mathbb{R} \to \mathbb{R}$ which are *measurable* with respect to $\mathcal{B}$, in a sense that we will define today. It turns out to come at no cost if instead of functions $f : \mathbb{R} \to \mathbb{R}$ we want to write our definitions and basic propositions in reference to functions $f : X \to Y$ where $(X, \mathfrak{X})$ and $(Y, \mathfrak{Y})$ are arbitrary measurable spaces.

For convenience, let $(X, \mathfrak{X}), (Y, \mathfrak{Y}), (Z, \mathfrak{Z})$ be measurable spaces throughout this subsection unless otherwise stated.

Def'n 2.8. **Measurable** Function

Let $f : X \to Y$. We say f is $\mathfrak{X} / \mathfrak{Y}$***-measurable*** if $f^{-1}(S) \in \mathfrak{X}$ for all $S \in \mathfrak{Y}$.

The preceding definition calls on the notion of preimage under f. Here are some useful properties of preimages.

Proposition 2.11.

Let $\{S_i\}_{i\in I}$ be a collection of subsets of Y and let $f : X \to Y$

(a) $f^{-1}\left(\bigcup_{i\in I} S_i\right) = \bigcup_{i\in I} f^{-1}(S_i)$.

(b) $f^{-1}\left(\bigcap_{i\in I} S_i\right) = \bigcap_{i\in I} f^{-1}(S_i)$.

(c) For any $S \subseteq Y$, $f^{-1}(Y \setminus S) = f^{-1}(Y) \setminus f^{-1}(S)$.

(d) Let $g : Y \to Z$. Then for all $T \subseteq Z$, then $(g \circ f)^{-1}(T) = f^{-1}\left(g^{-1}(T)\right)$.

See ProofWiki!

The most important instance of Def'n 2.8 is the one where $(Y, \mathfrak{Y})$ is the real line considered with its Borel σ-algebra, that is, $Y = \mathbb{R}$ and $\mathfrak{Y} = \mathcal{B}_{\mathbb{R}}$ (in what follows, we will write $\mathcal{B}_{\mathbb{R}}$ instead of $\mathcal{B}$ to avoid confusions).

Def'n 2.9. **Borel** Real-valued Functions

Given a measurable space $(X, \mathfrak{X})$, we will denote

$$\mathrm{Bor}(X, \mathbb{R}) = \left\{ f \in \mathbb{R}^X : f \text{ is } \mathfrak{X} / \mathcal{B}_{\mathbb{R}}\text{-measurable} \right\}.$$

We shall call the elements of $\mathrm{Bor}(X, \mathbb{R})$ ***Borel*** functions.

The notation introduced in Def'n 2.9 is bit imprecise, because it does not mention explicitly what σ-algebra of subsets of X is being considered. This is usually harmless, because it is clear from the context what is the σ-algebra $\mathfrak{X}$ we are working with. In a situation where there could be a possibility of confusion of what is $\mathfrak{X}$, we shall write $\mathrm{Bor}((X, \mathfrak{X}), \mathbb{R})$ instead.

Our main concern for the lecture is to examine what properties we can expect from the *space* of functions $\mathrm{Bor}(X, \mathbb{R})$. When calling $\mathrm{Bor}(X, \mathbb{R})$ space of functions, we are anticipating some good closure properties under various natural operations (addition, multiplication, …) with functions. And we will indeed be able to establish a number of such properties. But first we put into evidence some simple general tools which cna be used in order to study measurable functions. A couple of such *tools* are provided by the next two propositions.

Proposition 2.12. Composition of Measurable Functions Is Measurable

Let $f : X \to Y$ be an $\mathfrak{X} / \mathfrak{Y}$-measurable function and let $g : Y \to Z$ be an $\mathfrak{Y} / \mathfrak{Z}$-measurable function. Then $h = g \circ f$ is $\mathfrak{X} / \mathfrak{Z}$-measurable.

Proof. For all $C \subseteq Z$, we know that

$$h^{-1}(C) = (g \circ f)^{-1}(C) = f^{-1}\left(g^{-1}(C)\right)$$

by Proposition 2.11. This means

$$\begin{aligned} C \in \mathfrak{Z} &\implies g^{-1}(C) \in \mathfrak{Y} \\ &\implies f^{-1}\left(g^{-1}(C)\right) \in \mathfrak{X} \\ &\implies h^{-1}(C) \in \mathfrak{X}. \end{aligned}$$

Thus we conclude that h is $\mathfrak{X} / \mathfrak{Z}$-measurable, as required.

QED

Proposition 2.13.

Let $f : X \to Y$ and let $\mathcal{C} \subseteq \mathfrak{Y}$ be a collection of subsets of Y that generates $\mathfrak{Y}$: that is, $\sigma\text{-Alg}(\mathcal{C}) = \mathfrak{Y}$. If $f^{-1}(C) \in \mathfrak{X}$ for all $C \in \mathcal{C}$, then f is $\mathfrak{X} / \mathfrak{Z}$-measurable.

Proof. Let
$$\mathfrak{G} = \left\{S \subseteq Y : f^{-1}(S) \in \mathfrak{X}\right\}.$$

We have two claims.

- *Claim 1. $\mathfrak{G}$ is a σ-algebra of subsets of Y.*

  Proof. This is easily done by using basic properties of preimages. Suppose that $\mathcal{S}$ be a countable collection of sets from $\mathfrak{G}$, where we desire to check $\bigcup \mathcal{S} \in \mathfrak{G}$. The latter fact amounts to checking that $f^{-1}\left(\bigcup \mathcal{S}\right) \in \mathfrak{X}$. And indeed, using Proposition 2.11 we find
  $$f^{-1}\left(\bigcup_{n=1}^{\infty} S_n\right) = \bigcup_{n=1}^{\infty} f^{-1}(S_n). \tag{2.7}$$

  But for all $n \in \mathbb{N}, f^{-1}(S_n) \in \mathfrak{X}$, since $S_n \in \mathfrak{G}$. Hence the right-hand side of [2.7] is a countable union of sets from $\mathfrak{X}$, so is in $\mathfrak{X}$, since $\mathfrak{X}$ is a σ-algebra. Thus $\bigcup_{n=1}^{\infty} S_n \in \mathfrak{G}$, as required.

- *Claim 2. $\mathfrak{G} \supseteq \mathfrak{Y}$.*

  Proof. Since $f^{-1}(C) \in \mathfrak{X}$ for all $C \in \mathcal{C}$ by assumption, it follows that $\mathcal{C} \subseteq \mathfrak{G}$. This means $\mathfrak{G}$ is a σ-algebra containing $\mathcal{C}$, while, on the other hand, the hypothesis $\mathfrak{Y} = \sigma\text{-Alg}(\mathcal{C})$ tells us that $\mathfrak{Y}$ is the minimal σ-algebra containing $\mathcal{C}$. It follows that $\mathfrak{G} \supseteq \mathfrak{Y}$.

In conclusion, for every $S \in \mathfrak{Y}$, we have $S \in \mathfrak{G}$ by Claim 2, hence we have that $f^{-1}(S) \in \mathfrak{X}$. This amounts precisely to saying that f is $\mathfrak{X}/\mathfrak{Y}$-measurable.

QED

Corollary 2.13.1.

Let X, Y be metric spaces and let $f : X \to Y$ be continuous. Then f is $\mathcal{B}_X/\mathcal{B}_Y$-measurable.

Proof. Let $\mathcal{T}_X, \mathcal{T}_Y$ denote the collections of open sets of X, Y, respectively.

By Proposition 2.13, it suffices to show that
$$\forall G \in \mathcal{T}_Y \left[f^{-1}(G) \in \mathcal{B}_X\right].$$

But by a characterization of continuity,
$$G \in \mathcal{T}_Y \implies f^{-1}(G) \in \mathcal{T}_X \implies f^{-1}(G) \in \mathcal{B}_X$$

since $\mathcal{B}_X \supseteq \mathcal{T}_X$.

QED

We will also use a tool which deals specifically with functions that take values in a space $\mathbb{R}^n$.

Proposition 2.14.

Let $f = (f_1, \ldots, f_n) : X \to \mathbb{R}^n$ (that is, $f_1, \ldots, f_n : X \to \mathbb{R}$ are the *component functions* of f). The following are equivalent.

(a) f is $\mathfrak{X}/\mathcal{B}_{\mathbb{R}^n}$-measurable.

(b) Each f_j is $\mathfrak{X}/\mathcal{B}_{\mathbb{R}}$-measurable (i.e. $f_1, \ldots, f_n \in \text{Bor}(X, \mathbb{R})$).

Proof. Fix $j \in \{1, \ldots, n\}$, for which we will prove that $f_j \in \text{Bor}(X, \mathbb{R})$. Let $P_j : \mathbb{R}^n \to \mathbb{R}$ be the jth projection map, defined by
$$P(t_1, \ldots, t_n) = t_j$$

for all $(t_1, \ldots, t_n) \in \mathbb{R}^n$. Then P_j is continuous, so $\mathcal{B}_{\mathbb{R}^n}/\mathcal{B}_{\mathbb{R}}$-measurable. It follows that $f_j = P_j \circ f$ is $\mathfrak{X}/\mathcal{B}_{\mathbb{R}}$-measurable.

Conversely, suppose that $f_1, \ldots, f_n$ are $\mathfrak{X}/\mathcal{B}_{\mathbb{R}}$-measurable, where we want to prove that f is $\mathcal{B}_{\mathbb{R}^n}/\mathcal{B}_{\mathbb{R}}$-measurable. We will do the proof by using Proposition 2.13 in connection to the collection $\mathcal{C}$ of *open cubes* in $\mathbb{R}^n$, defined as follows:
$$\mathcal{C} = \{(a_1 - r, a_1 + r) \times \cdots \times (a_n - r, a_n + r) : a_1, \ldots, a_n \in \mathbb{R}, r \in (0, \infty)\}.$$

In connection this collection of subsets of $\mathbb{R}^n$, we make two claims.

- *Claim 1.* σ-Alg$(\mathcal{C}) = \mathcal{B}_{\mathbb{R}^n}$.

  Proof. Note that $\mathcal{C}$ is precisely the collection of open balls with respect to $\|\cdot\|_\infty$. But by recalling the fact that every norm on $\mathbb{R}^n$ are equivalent and that the collection of open balls generate the usual topology (i.e. $\mathcal{C}$ is a basis for the usual topology on $\mathbb{R}^n$), we arrive to the conclusion that σ-Alg$(\mathcal{C}) = \mathcal{B}_{\mathbb{R}^n}$.

- *Claim 2. For all $C \in \mathcal{C}$, $f^{-1}(C) \in \mathfrak{X}$.*

  Proof. Let $C = (a_1 - r, a_1 + r) \times \cdots \times (a_n - r, a_n + r) \in \mathcal{C}$. Since each f_j is $\mathfrak{X} / \mathcal{B}_{\mathbb{R}}$-measurable, it follows that $f_j^{-1}\left((a_j - r, a_j + r)\right) \in \mathfrak{X}$. This means

  $$f^{-1}(S) = \bigcap_{j=1}^{n} f_j^{-1}\left((a_j - r, a_j + r)\right)$$

  is also in $\mathfrak{X}$.

By invoking Proposition 2.13, we are done.

QED

We now arrive to the main point of this subsection, concerning the closure of Bor$(X, \mathbb{R})$ under various algebraic operations that can be performed with real-valued functions. Here are some *pointwise* operations we can perform on $f, g : X \to \mathbb{R}$.

(a) Given $\alpha, \beta \in \mathbb{R}$, consider $\alpha f + \beta g$, defined by $(\alpha f + \beta g)(x) = \alpha f(x) + \beta g(x)$ for all $x \in \mathbb{R}$. *linear combination*

(b) Consider fg, defined by $(fg)(x) = f(x) g(x)$ for all $x \in \mathbb{R}$. *product*

(c) We define $f \vee g, f \wedge g$ by

$$(f \vee g)(x) = \max(f(x), g(x))$$

and by

$$(f \wedge g)(x) = \min(f(x), g(x))$$

for all $x \in \mathbb{R}$. *maximum and minimum*

It turns out that Bor$(X, \mathbb{R})$ is closed under all these operations.

Proposition 2.15.

Let $f, g \in$ Bor$(X, \mathbb{R})$ and let $\alpha, \beta \in \mathbb{R}$. Then $fg, f \vee g, f \wedge g, \alpha f + \beta g \in$ Bor$(X, \mathbb{R})$.

Proof. We verify $f \vee g \in$ Bor$(X, \mathbb{R})$ only.

Note that $F : X \to \mathbb{R}^2$ by $F = (f, g)$ is $\mathfrak{X} / \mathcal{B}_{\mathbb{R}^2}$-measurable by Proposition 2.14. Moreover, $\max : \mathbb{R}^2 \to \mathbb{R}$ is continous, so is $\mathcal{B}_{\mathbb{R}^2} / \mathcal{B}_{\mathbb{R}}$-measurable. Thus $f \vee g = \max \circ F$ is $\mathfrak{X} / \mathcal{B}_{\mathbb{R}}$-measurable.

QED

Proposition 2.15 tells us that Bor$(X, \mathbb{R})$ is a (unital) algebra of functions and is also a lattice of functions.

Exercise 2.5.

(a) Let $\mathcal{C} = \{(-\infty, b] : b \in \mathbb{Q}\}$. Prove that σ-Alg$(\mathcal{C}) = \mathcal{B}_{\mathbb{R}}$.

(b) Let $f : X \to \mathbb{R}$. Suppose for every $b \in \mathbb{Q}$, $\{x \in X : f(x) \le b\}$ is in $\mathfrak{X}$. Prove $f \in$ Bor$(X, \mathbb{R})$.

tl;dr

4. Convergence Properties of Bor$(X, \mathbb{R})$

We now turn to properties of Bor$(X, \mathbb{R})$ which have to do with *pointwise convergnece* of sequences of functions. These will turn out to be as good as well: the space Bor$(X, \mathbb{R})$ is closed under pointwise convergence of sequences of functions, and we also have some variations of this fact where instead of a limit we play with inf and sup constructions.

The main result of this subsection is in Proposition 2.16, which can be viewed as yet another tool for establishing measurability of a function. We will not give the proof of Proposition 2.16 right away; it will be a particular case of the more general Proposition 2.18.

Proposition 2.16.

Let $(X, \mathcal{A})$ be a measurable space and let $f : X \to \mathbb{R}$. If there exists $(f_n)_{n=1}^{\infty} \in \text{Bor}(X, \mathbb{R})^{\mathbb{N}}$ such that $\lim_{n\to\infty} f_n = f$ pointwise, then $f \in \text{Bor}(X, \mathbb{R})$.

This Is a Particular Instance of Proposition 2.18

Example 2.6.

Here is an example of how Proposition 2.16 can be used. It goes in the framework of the measurable space $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ – the space $\text{Bor}(X, \mathbb{R})$ of this example is thus $\text{Bor}(\mathbb{R}, \mathbb{R})$.

We consider the following statement: *for all differentiable* $f : \mathbb{R} \to \mathbb{R}$, $f' \in \text{Bor}(\mathbb{R}, \mathbb{R})$.

From calculus, we know that derivatives can sometimes be *nasty* (e.g. a derivative need not be continuous). The above statement says that f' cannot be *that* nasty as to not be a Borel function.

In order to prove that $f' \in \text{Bor}(\mathbb{R}, \mathbb{R})$, we write it as a pointwise limit

$$f'(t) = \lim_{n\to\infty} f_n(t)$$

where $f_n : \mathbb{R} \to \mathbb{R}$ is defined by putting

$$f_n(t) = \frac{f\left(t + \frac{1}{n}\right) - f(t)}{\frac{1}{n}}$$

for all $t \in \mathbb{R}$. Then note that each f_n is continuous, so $f_n \in \text{Bor}(\mathbb{R}, \mathbb{R})$. Thus $f' \in \text{Bor}(\mathbb{R}, \mathbb{R})$ as a pointwise limit of f_n.

Now let us start to work towards proving that Proposition 2.16 is indeed available. We will actually establish a more refined version of this statement, which instead of lim calls in for a lim sup or a lim inf. On the way towards that, we first get a version which is merely using sup or inf.

Proposition 2.17.

Let $(X, \mathcal{A})$ be a measurable space.

(a) Let $f : X \to \mathbb{R}$. If there exists $(f_n)_{n=1}^{\infty} \in \text{Bor}(X, \mathbb{R})^{\mathbb{N}}$ such that

$$\sup_{n\in\mathbb{N}} f_n(x) = f(x)$$

for all $x \in X$, then $f \in \text{Bor}(X, \mathbb{R})$.

(b) Let $g : X \to \mathbb{R}$. If there exists $(g_n)_{n=1}^{\infty} \in \text{Bor}(X, \mathbb{R})^{\mathbb{N}}$ such that

$$\inf_{n\in\mathbb{N}} g_n(x) = g(x)$$

for all $x \in X$, then $g \in \text{Bor}(X, \mathbb{R})$.

Proof. We make use of the criterion

$$\forall t \in \mathbb{R} \left[f^{-1}((-\infty, t]) \in \mathcal{A} \right] \implies f \in \text{Bor}(X, \mathbb{R})$$

provided in Exercise 2.5.[1] To that end, fix $t \in \mathbb{R}$. In connection to t, for every $x \in X$ we have

$$f(x) \le t \iff \sup_{n\in\mathbb{N}} f_n(x) \le t \iff \forall n \in \mathbb{N} \left[f_n(x) \le t \right].$$

This implies we have

$$f^{-1}((-\infty, t]) = \bigcap_{n=1}^{\infty} f_n^{-1}((-\infty, t]). \qquad [2.8]$$

For every $n \in \mathbb{N}$, $f_n^{-1}((-\infty, t]) \in \mathcal{A}$, since it is the preimage of the interval $(-\infty, t] \in \mathcal{B}_{\mathbb{R}}$ by the $\mathcal{A} / \mathcal{B}_{\mathbb{R}}$-measurable functoin f_n. This means that $f^{-1}((-\infty, t])$ is written as a countable intersection of sets from $\mathcal{A}$ in [2.8]. It follows that $f^{-1}((-\infty, t]) \in \mathcal{A}$, as required. This completes (a).

For (b), it suffices to note that

$$g(x) = -\sup_{n\in\mathbb{N}} -g_n(x)$$

for all $x \in X$, and use the result that $\mathrm{Bor}(X, \mathbb{R})$ is closed under linear combinations.

[1] In fact, it suffices to check $f^{-1}((-\infty, t]) \in \mathcal{A}$ for all $t \in \mathbb{Q}$ only.

QED

The next (general) version of Proposition 2.16 will rely on the notions of lim sup and lim inf for a sequence of real numbers – let us recall what those are.

Let $(t_n)_{n=1}^{\infty} \in \mathbb{R}^{\mathbb{N}}$, and let us consider the set of all limit points of this sequence. That is, we look $P \subseteq \mathbb{R} \cup \{\pm\infty\}$ defined as follows

$$P = \left\{ l \in \mathbb{R} \cup \{\pm\infty\} : \exists n_1, n_2, \ldots \in \mathbb{N}\left[n_1 < n_2 < \cdots, \lim_{k\to\infty} t_{n_k} = l \right] \right\}.$$

It can be proved that P has a largest element and a smallest element. We define

$$\limsup_{n\to\infty} t_n = \max(P), \liminf_{n\to\infty} t_n = \min(P).$$

An alternative way to reach the quantity $\max(P)$ is described as follows. Consider

$$\beta_k = \sup_{n\geq k} t_n$$

for all $k \in \mathbb{N}$. Then $(\beta_k)_{k=1}^{\infty}$ is a decreasing sequence of real numbers, and it can be proved that

$$\max(P) = \lim_{k\to\infty} \beta_k = \inf_{k\in\mathbb{N}} \beta_k.$$

Therefore, we obtain

$$\limsup_{n\to\infty} t_n = \inf_{k\geq 1} \sup_{n\geq qk} t_n. \quad [2.9]$$

Similarly

$$\liminf_{n\to\infty} t_n = \sup_{k\geq 1} \inf_{n\geq k} (t_n). \quad [2.10]$$

Proposition 2.18.

Let $(X, \mathcal{A})$ be a measurable space.

(a) Let $f : X \to \mathbb{R}$. If there exists $(f_n)_{n=1}^{\infty} \in \mathrm{Bor}(X, \mathbb{R})^{\mathbb{N}}$ such that

$$\limsup_{n\to\infty} f_n(x) = f(x)$$

for all $x \in X$, then $f \in \mathrm{Bor}(X, \mathbb{R})$.

(b) Let $g : X \to \mathbb{R}$. If there exists $(g_n)_{n=1}^{\infty} \in \mathrm{Bor}(X, \mathbb{R})^{\mathbb{N}}$ such that

$$\liminf_{n\to\infty} g_n(x) = g(x)$$

for all $x \in X$, then $g \in \mathrm{Bor}(X, \mathbb{R})$.

Proof. We start by noticing that, for every $x \in X$, the sequence $(f_n(x))_{n=1}^{\infty}$ is bounded above. Indeed, if that was not the case, then we would be able to extract a subsequence of $(f_n(x))_{n=1}^{\infty}$ which converges to ∞, and this would imply $\limsup_{n\to\infty} f_n(x) = \infty$, in contradiction with the given hypothesis that $\limsup_{n\to\infty} f_n(x) = f(x) \in \mathbb{R}$. So it makes sense to define $h_1 : X \to \mathbb{R}$ by putting

$$h_1(x) = \sup_{n\in\mathbb{N}} f_n(x)$$

for all $x \in X$. Moreover, Proposition 2.17 applies here, to give us that $h_1 \in \mathrm{Bor}(X, \mathbb{R})$.

We can similarly define, for all $k \geq 2$,

$$h_x(x) = \sup_{n \geq k} f_n(x)$$

for all $x \in X$. Again, each $h_k \in \mathrm{Bor}(X, \mathbb{R})$.

We note that, for all $x \in X$, the sequence $(h_k(x))_{k=1}^{\infty}$ is decreasing, with

$$\inf_{k \in \mathbb{N}} h_k(x) = \inf_{k \geq 1} \sup_{n \geq k} f_n(x) = \limsup_{n \to \infty} f_n(x) = f(x),$$

where the second equality sign we invoked [2.9].

This means f is presented as the (pointwise) infimum of the sequence of functions $(n_k)_{k=1}^{\infty}$. Since $h_k \in \mathrm{Bor}(X, \mathbb{R})$ for all $k \geq 1$, we can invoke Proposition 2.17 to conclude that $f \in \mathrm{Bor}(X, \mathbb{R})$. This marks the end of (a).

For (b), we once again use the trick $f = -g, f_n = -g_n$ for all $N \in \mathbb{N}$.

QED

At this moment, we have also settled the Proposition 2.16 stated at the beginning of this subsection. Indeed, let f and $(f_n)_{n=1}^{\infty}$ be as in Proposition 2.16. Then for every $x \in X, f(x)$ is the unique limit poitn of $(f_n(x))_{n=1}^{\infty}$, so that

$$\limsup_{n \to \infty} f_n(x) = \liminf_{n \to \infty} f_n(x) = f(x).$$

By invoking Proposition 2.18, we conclude that $f \in \mathrm{Bor}(X, \mathbb{R})$, as required.

The statement of Proposition 2.18 was arranged in such a way that all the lim sup's and lim inf's considered there are forced to be finite. This is because there is a rescribed function f, taking finite values, which is arranged to be the lim sup or lim inf in question. Let us also have a look at what happens when we don't prescribe such an f.

So let $(X, \mathcal{A})$ be a measurable space, and let $(f_n)_{n=1}^{\infty} \in \mathrm{Bor}(X, \mathbb{R})^{\mathbb{N}}$. Say we want to look at the lim sup of $(f_n)_{n=1}^{\infty}$. It is obvious how we should define it, only that now we have to deal with the possibly nonempty sets

$$A_+ = \left\{ x \in X : \limsup_{n \to \infty} f_n(x) = 0 \right\}, A_- = \left\{ x \in X : \limsup_{n \to \infty} f_n(x) = -\infty \right\}.$$

Exercise 2.7.

Prove that $A_+, A_- \in \mathcal{A}$.

tl;dr

Suppose next that we have decided to ignore what happens for the ponits $x \in A_+ \cup A_-$, and we go ahead to define $f : X \to \mathbb{R}$ by putting

$$f(x) = \begin{cases} 0 & \text{if } x \in A_+ \cup A_- \\ \limsup_{n \to \infty} f_n(x) & \text{otherwise} \end{cases}$$

for all $x \in X$.

Exercise 2.8.

Prove that $f \in \mathrm{Bor}(X, \mathbb{R})$.

tl;dr

Upon solving it, we will thus have proved a result about $\mathrm{Bor}(X, \mathbb{R})$ *being well-behaved under* lim sup, where we don't prescribe a target function f which is to appear as lim sup.

Let $(X, \mathcal{A})$ be a measurable space throughout this subsection.

Def'n 2.10. **Simple** Function
We say $f \in \operatorname{Bor}(X, \mathbb{R})$ is ***simple*** when it takes finitely many values (i.e. $|f(X)| < |\mathbb{N}|$). We write $\operatorname{Bor}_s(X, \mathbb{R})$ to denote the set of simple functions in $\operatorname{Bor}(X, \mathbb{R})$.

One has a natural way of writing a general function $f \in \operatorname{Bor}(X, \mathbb{R})$ as a pointwise limit of functions from $\operatorname{Bor}_s(X, \mathbb{R})$. For convenience, we will focus here on the case when f only takes values in $[0, \infty)$. The procedure for finding functions in $\operatorname{Bor}_s(X, \mathbb{R})$ which approximate f can be thought of as a special way of *binning the values of f*.

More precisely, suppose we are given a function $f \in \operatorname{Bor}(X, \mathbb{R})$ such that $f(x) \geq 0$ for all $x \in X$. For all $n \in \mathbb{N}$, let us use the name *n-th binning* of f for $f_n : X \to \mathbb{R}$ defined as follows:

$$f_n(x) = \begin{cases} n & \text{if } f(x) \geq n \\ \frac{k-1}{2^n} & \text{otherwise} \end{cases}$$

for all $n \in \mathbb{N}$, where, in case $f(x) < n$, k is the unique element of $\{1, \ldots, n2^n\}$ such that $\frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n}$.

Proposition 2.19.
Let $f \in \operatorname{Bor}(X, \mathbb{R})$ and let $(f_n)_{n=1}^{\infty}$ be defined as above. Then
(a) for all $n \in \mathbb{N}$, $f_n \in \operatorname{Bor}_s(X, \mathbb{R})$;
(b) for all $n \in \mathbb{N}$, $f_n \leq f_{n+1}$;[1]
(c) for all $x \in X$, $f_n(x) \nearrow f(x)$.[2]

[1] We write $f \leq g$ for $f, g : X \to \mathbb{R}$ whenever $f(x) \leq g(x)$ for all $x \in \mathbb{R}$.
[2] We write $f_n(x) \nearrow f(x)$ whenever $(f_n(x))_{n=1}^{\infty}$ is an *increasing sequence that converges to* $f(x)$.

Exercise

A salient detail which is sure to appear during the verifications is this:

$$\forall x \in X \exists n_0 \in \mathbb{N}\, \forall n \geq n_0 \left[f(x) - \frac{1}{2^n} \leq f_n(x) \leq f(x) \right],$$

from where (c) of Proposition 2.19 follows.

III. Lebesgue Integration

1. Lebesgue Integration of Nonnegative Simple Borel Functions

The preceding section devoted to the space of functions $\mathrm{Bor}(X, \mathbb{R})$ have prepared us for taking on the notion of *integral* for functions in that space.

The construction of the Lebesgure works like a charm in the general framework of a *measure space*, a triple $(X, \mathcal{A}, \mu)$ where X is a nonempty set, $\mathcal{A}$ is a σ-algebra of subsets of X, and $\mu : \mathcal{A} \to [0, \infty]$ is a positive measure. Our primary measure space of interest is $(\mathbb{R}, \mathcal{B}, \lambda)$, where $\mathcal{B}$ is the Borel σ-algebra of subsets of $\mathbb{R}$ and $\lambda : \mathcal{B} \to [0, \infty]$ is the Lebesgue measure. But it comes at no cost, and is in fact more transparent, if we develop the Lebesgure integral for a general measure space $(X, \mathcal{A}, \mu)$.

So for the next few subsectsions we will stick to the $(X, \mathcal{A}, \mu)$ framework. Thus, fix a measure space $(X, \mathcal{A}, \mu)$ throughout.

Recall that μ is *finite sub-additive*:

$$\forall A_1, \ldots, A_k \in \mathcal{A}\left[\mu\left(\bigcup_{j=1}^{k} A_j\right) \leq \sum_{j=1}^{k} \mu\left(A_j\right)\right].$$

We would like to extend the sub-additivity to countable unions, and also record the good behavior of μ with respect to increasing or decreasing chains of sets in $\mathcal{A}$.

Proposition 3.1.

(a) Let $(A_n)_{n=1}^{\infty} \in \mathcal{A}^{\mathbb{N}}$ be an increasing chain. Then

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu\left(A_n\right) \in [0, \infty]. \qquad \textit{continuity along increasing chain}$$

(b) Let $(B_n)_{n=1}^{\infty} \in \mathcal{A}^{\mathbb{N}}$ be a decreasing chain such that $\mu(B_1) < \infty$. Then

$$\mu\left(\bigcap_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} \mu\left(B_n\right) \in [0, \infty). \qquad \textit{continuity along decreasing chain}$$

(c) Let $\mathcal{C} \subseteq \mathcal{A}$ be countable. Then

$$\mu\left(\bigcup \mathcal{C}\right) = \sum_{C\in\mathcal{C}} \mu\left(C\right). \qquad \textit{countable sub-additivity}$$

Proof.

(a) Let

$$D_n = A_n \setminus A_{n-1}$$

(with $A_0 = \varnothing$) for all $n \in \mathbb{N}$. Then for all $n \in \mathbb{N}$, $D_n \in \mathcal{A}$, as $\mathcal{A}$ is closed under set-difference. It is moreover immediate that D_n's are pairwise disjoint. Hence by the σ-additivity of μ, we have

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

But the (finite) additivity of μ can be invoked as well, to obtain

$$\forall k \in \mathbb{N}\left[\mu\left(\bigcup_{n=1}^{k} D_n\right) = \sum_{n=1}^{k} \mu\left(D_n\right)\right].$$

By definition, $A_k = \bigcup_{n=1}^{k} D_n$ for all $k \in \mathbb{N}$, so that $\bigcup_{n=1}^{\infty} D_n = \bigcup_{n=1}^{\infty} A_n$. We can therefore write

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

(b) The idea of the proof is to *subtract all the B_n's out of B_1*. We will be able to handle the behavior of μ under these set-differences, due to the hypothesis that each $\mu(B_n)$ is a finite number (so that we can safely perform differences $\mu(B_1) - \mu(B_n)$ without running into $\infty - \infty$).

To be precise, for all $n \in \mathbb{N}$, let

$$A_n = B_1 \setminus B_n.$$

It is immediate that $(A_n)_{n=1}^{\infty}$ is an increasing chain of sets in $\mathcal{A}$, with

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} (B_1 \setminus B_n) = B_1 \setminus \left(\bigcap_{n=1}^{\infty} B_n\right).$$

Hence the continuity along increasing chain tells us that

$$\lim_{n\to\infty} \mu(A_n) = \mu\left(B_1 \setminus \left(\bigcap_{n=1}^{\infty} B_n\right)\right). \tag{3.1}$$

For every $n \in \mathbb{N}$, the way how A_n was defined implies that we have $A_n \cup B_n = B_1$ with $A_n \cap B_n = \varnothing$. This in turn implies $\mu(A_n) + \mu(B - n) = \mu(B_1)$, leading to the formula

$$\mu(A_n) = \mu(B_1) - \mu(B_n).$$

We again remark that the above quantity is well-defined since we assumed $\mu(B_1) < \infty$. Likewise, $\mu\left(B_1 \setminus \left(\bigcap_{n=1}^{\infty} B_n\right)\right) = \mu(B_1) - \mu\left(\bigcap_{n=1}^{\infty} B_n\right)$, which makes the limit from [3.1] take the form

$$\lim_{n\to\infty} \mu(B_1) - \mu(B_n) = \mu(B_1) - \mu\left(\bigcap_{n=1}^{\infty} B_n\right). \tag{3.2}$$

From [3.2], some straightforward work with finite limits of convergent sequencces leads to the desired formula.

(c) Since we know μ is finite-additive, we may assume $\mathcal{C}$ is countable; say $\mathcal{C} = \{C_n\}_{n=1}^{\infty}$.

Let us denote $C = \bigcup_{n=1}^{\infty} C_n$ and $\gamma = \sum_{n=1}^{\infty} \mu(C_n)$. Then it suffices to check $\mu(C) \leq \gamma$. In order to do so, we will work with

$$A_n = \bigcup_{k=1}^{n} C_k,$$

defined for all $n \in \mathbb{N}$. It is immediate that $(A_n)_{n=1}^{\infty}$ is an increasing chain of sets from $\mathcal{A}$, with $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} C_n = C$. Hence

$$\mu(C) = \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n), \tag{3.3}$$

where at the second equality sign we invoked the continuity of μ along increasing chain. In view of [3.3], we see that the desired inequality $\mu(C) \leq \gamma$ will follow if we can verify that $\mu(A_n) \leq \gamma$ for all $n \in \mathbb{N}$.

Indeed, for any $n \in \mathbb{N}$, we can write

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

Thus we see that $\mu(C) \leq \gamma$, as required.

QED

We are now good to go on integration. Since the Lebesgue integral is such an important notion, we will take time to build it, and will introduce it in the guise of a progressive construction of functionals defined on some subsets of the space of functions $\mathrm{Bor}(X, \mathbb{R})$. For this time, we consider only those are *simple* and *nonnegative*.

Notation 3.1. $\mathrm{Bor}^{+}(X,\mathbb{R}), \mathrm{Bor}_{\mathrm{s}}^{+}(X,\mathbb{R})$
Let us denote

$$\mathrm{Bor}^{+}(X,\mathbb{R}) = \{f \in \mathrm{Bor}(X,\mathbb{R}) : \forall x \in X [f(x) \geq 0]\},$$

the set of nonnegative Borel functions. We then denote

$$\mathrm{Bor}_{\mathrm{s}}^{+}(X,\mathbb{R}) = \mathrm{Bor}_{\mathrm{s}}(X,\mathbb{R}) \cap \mathrm{Bor}^{+}(X,\mathbb{R}),$$

the set of simple nonnegative Borel functions.

We found that $\mathrm{Bor}(X,\mathbb{R})$ is a unital algebra and a lattice. It is immediate that $\mathrm{Bor}_{\mathrm{s}}(X,\mathbb{R})$ is both a unital subalgebra and a sublattice of $\mathrm{Bor}(X,\mathbb{R})$.

The subset $\mathrm{Bor}^{+}(X,\mathbb{R})$ is not a subspace of $\mathrm{Bor}(X,\mathbb{R})$, but it is obvious that we nevertheless have

$$\alpha f + \beta g \in \mathrm{Bor}^{+}(X,\mathbb{R})$$

for all $f, g \in \mathrm{Bor}^{+}(X,\mathbb{R}), \alpha, \beta \in [0,\infty)$ and also that $\mathrm{Bor}^{+}(X,\mathbb{R})$ is closed under multiplication and is a sublattice of $\mathrm{Bor}(X,\mathbb{R})$. These closure properties of $\mathrm{Bor}^{+}(X,\mathbb{R})$ are passed alongg to $\mathrm{Bor}_{\mathrm{s}}^{+}(X,\mathbb{R})$ as well.

As a matter of notation, let us use the following to denote the zero vector of $\mathrm{Bor}(X,\mathbb{R})$.

Notation 3.2. $\underline{0}$
We will write $\underline{0}$ to denote the zero vector of $\mathrm{Bor}(X,\mathbb{R})$. That is, $\underline{0} : X \to \mathbb{R}$ is defined by

$$\underline{0}(x) = 0$$

for all $x \in X$.

We are now ready to introduce the map $\mathrm{L}_{\mathrm{s}}^{+} : \mathrm{Bor}_{\mathrm{s}}^{+}(X,\mathbb{R}) \to [0,\infty]$ which is our first stint at integration.

Notation 3.3. $\mathrm{L}_{\mathrm{s}}^{+}$
We define $\mathrm{L}_{\mathrm{s}}^{+} : \mathrm{Bor}_{\mathrm{s}}^{+}(X,\mathbb{R}) \to [0,\infty]$ as follows. Given any $f \in \mathrm{Bor}_{\mathrm{s}}^{+}(X,\mathbb{R})$, we consider $\alpha_1, \ldots, \alpha_p$, with $\alpha_1 < \cdots < \alpha_p$, the complete list of values assumed by f and

$$A_i = f^{-1}(\alpha_i)$$

for all $i \in \{1, \ldots, p\}$. We then put

$$\mathrm{L}_{\mathrm{s}}^{+}(f) = \sum_{i=1}^{p} \alpha_i \mu(A_i) \in [0,\infty]. \quad [3.4]$$

In order for the formula [3.4] used to define $\mathrm{L}_{\mathrm{s}}^{+}(f)$ makes sense, we need to make some clarifications.

(a) Each A_i is the preimage of $\{\alpha_i\}$, a Borel set, under a $\mathcal{A}/\mathcal{B}_{\mathbb{R}}$-measurable function, so $\mu(A_i)$ is well-defined.
(b) If, for some $i \in \{1, \ldots, p\}$, $\alpha_i > 0$ and $\mu(A_i) = \infty$, then we have $\alpha_i \mu(A_i) = \infty$, leading to the conclusion that $\mathrm{L}_{\mathrm{s}}^{+}(f) = \infty$.
(c) In case $\alpha_i = 0, \mu(A_i) = \infty$ for some $i \in \{1, \ldots, p\}$ (in fact, this can only happen for $i = 1$), we use the following convention:

$$0 \cdot \infty = 0.$$

A clear test-case for the above discussion:

$$\mathrm{L}_{\mathrm{s}}^{+}(\underline{0}) = 0\mu(X) = 0,$$

holding no matter what $\mu(X)$ is.

Def'n 3.4. **Indicator Function** of a Subset

Let $A \subseteq X$. We define the ***indicator function*** of A, denoted as χ_A, by

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

for all $x \in X$.

Example 3.1. Integration of Indicator Functions

Let $A \in \mathcal{A}$. Then it is immediate that $\chi_A \in \mathrm{Bor}_s^+(X, \mathbb{R})$, since a preimage $\chi_A^{-1}(S)$ is one of $\varnothing, A$ for all $S \subseteq \mathbb{R}$. In this special case, [3.4] gives that

$$\mathrm{L}_s^+\left(\chi_A\right) = \mu(A), \tag{3.5}$$

the Lebesgue measure of A.

In some sense, one can say that the definition of L_s^+ *goes by pushing to linear combinations* in the above equation [3.5]. More precisely, given $f \in \mathrm{Bor}_s^+(X, \mathbb{R})$ which takes values $\alpha_1, \ldots, \alpha_p$, it is then immediately checked that we have

$$f = \sum_{i=1}^{p} \alpha_i \chi_{f^{-1}(\{\alpha_i\})}.$$

Hence the formula [2.4] defining $\mathrm{L}_s^+(f)$ can be read as

$$\mathrm{L}_s^+(f) = \sum_{i=1}^{p} \alpha_i \, \mathrm{L}_s^+\left(\chi_{f^{-1}(\{\alpha_i\})}\right).$$

Here are some more basic properties of L_s^+.

Exercise 3.2. Positive Homogeneity of L_s^+

Prove that

$$\mathrm{L}_s^+(\alpha f) = \alpha \, \mathrm{L}_s^+(f)$$

for every $\alpha \in [0, \infty), f \in \mathrm{Bor}_s^+(X, \mathbb{R})$.

We next want to show that L_s^+ has nice behavior under addition. To that end, it is useful to note, in the next lemma, a slight rephrasing of how L_s^+ was defined.

Lemma 3.2.

Let $f \in \mathrm{Bor}_s^+(X, \mathbb{R})$. If there exist $\gamma_1, \ldots, \gamma_r$ such that $\left\{f^{-1}\left(\gamma_j\right)\right\}_{j=1}^{r}$ is a partition of X, then

$$\mathrm{L}_s^+(f) = \sum_{j=1}^{r} \gamma_j \mu\left(f^{-1}\left(\gamma_j\right)\right).$$

Proof Sketch. Let $C_j = f^{-1}\left(\gamma_j\right)$ for all $J \in \{1, \ldots, r\}$.

We have to show that

$$\sum_{j=1}^{r} \gamma_j \mu\left(f^{-1}\left(\gamma_j\right)\right) = \sum_{i=1}^{p} \alpha_i \mu\left(A_i\right), \tag{3.6}$$

where $\alpha_1 < \cdots < \alpha_p$ are the distinct values taken by f and $A_i = f^{-1}(\alpha_i)$ for all $i \in \{1, \ldots, p\}$.

When we compare the two sides of the above equation [2.6], we see that they only differ in the respect that, some of $f^{-1}\left(\gamma_j\right)$'s may be empty and the numbers $\gamma_1, \ldots, \gamma_r$ may not be distinct. So then what we have to do is to remove the $f^{-1}\left(\gamma_j\right)$'s that are empty and put together the γ_j's that are equal to each other – when that is done, the left-hand side of [2.6] will be converted precisely into the right-hand side of the equation, with every A_i appearing as a disjoint union of $f^{-1}\left(\gamma_j\right)$'s.

QED

Proposition 3.3. Additivity of L_s^+

For every $f, g \in \mathrm{Bor}_\mathrm{s}^+(X, \mathbb{R})$,

$$\mathrm{L}_\mathrm{s}^+(f+g) = \mathrm{L}_\mathrm{s}^+(f) + \mathrm{L}_\mathrm{s}^+(g).$$

Proof. Let $\alpha_1, \ldots, \alpha_p$ be the values assumed by f, with $\alpha_1 < \cdots < \alpha_p$, and let $A_i = f^{-1}(\alpha_i)$ for all $i \in \{1, \ldots, p\}$. Let $\beta_1, \ldots, \beta_q \in \mathbb{R}, B_1, \ldots, B_q$ be defined in an analogous way. Consider

$$C_{i,j} = A_i \cap B_j$$

and

$$\gamma_{i,j} = \alpha_i + \beta_j$$

for all $i \in \{1, \ldots, p\}, j \in \{1, \ldots, q\}$. Then $C_{i,j} = (f+g)^{-1}\left(\gamma_{i,j}\right)$ for all $1, \ldots, p$, So by Lemma 3.2,

$$\mathrm{L}_\mathrm{s}^+(f+g) = \sum_{1 \leq i \leq p, 1 \leq j \leq q} \gamma_{i,j} \mu\left(C_{i,j}\right).$$

Upon replacing the $\gamma_{i,j}$ and $C_{i,j}$ from the formulas defining them, we get

$$\mathrm{L}_\mathrm{s}^+(f+g) = \sum_{1 \leq i \leq p, 1 \leq j \leq q} \left(\alpha_i + \beta_j\right) \mu\left(A_i \cap B_j\right) = \sum_{i=1}^{p} \alpha_i \sum_{j=1}^{q} \mu\left(A_i \cap B_j\right) + \sum_{j=1}^{q} \beta_j \sum_{i=1}^{p} \mu\left(A_i \cap B_j\right). \tag{3.7}$$

But for every $i \in \{1 \ldots, p\}$, $A_i \cap B_1, \ldots, A_i \cap B_q$ form a partition of A_i, so by the additivity of μ,

$$\sum_{j=1}^{q} \mu\left(A_i \cap B_j\right) = \mu\left(A_i\right).$$

In a similar manner, we see that

$$\sum_{i=1}^{p} \mu\left(A_i \cap B_j\right) = \mu\left(B_j\right)$$

for all $j \in \{1, \ldots, q\}$. Thus [3.7] becomes

$$\mathrm{L}_\mathrm{s}^+(f+g) = \sum_{i=1}^{p} \alpha_i \mu\left(A_i\right) + \sum_{j=1}^{q} \beta_j \mu\left(B_j\right) = \mathrm{L}_\mathrm{s}^+(f) + \mathrm{L}_\mathrm{s}^+(g),$$

which is what we intended to show.

QED

There is a natural partial order $\leq$ on $\mathrm{Bor}(X, \mathbb{R})$ introduced by declaring that, for every $f, g \in \mathrm{Bor}(X, \mathbb{R})$, we have

$$f \leq g \iff \forall x \in X\,[f(x) \leq g(x)].$$

The reason for mentioning $\leq$ is that we observe the following immediate consequence of Proposition 3.3.

Corollary 3.3.1. L_s^+ Is Increasing

Let $f, g \in \mathrm{Bor}^+(X, \mathbb{R})$. If $f \leq g$, then $\mathrm{L}_\mathrm{s}^+(f) \leq \mathrm{L}_\mathrm{s}^+(g)$.

Proof. Let $h = g - f$. Then $h \in \mathrm{Bor}_\mathrm{s}(X, \mathbb{R})$ since $\mathrm{Bor}_\mathrm{s}(X, \mathbb{R})$ is a vector subspace of $\mathrm{Bor}(X, \mathbb{R})$, and $h \in \mathrm{Bor}^+(X, \mathbb{R})$ due to the hypothesis that $f \leq g$. This means $h \in \mathrm{Bor}_\mathrm{s}^+(X, \mathbb{R})$ such that $h + f = g$. By applying Proposition 3.3 to this situation, we find that

$$\mathrm{L}_\mathrm{s}^+(g) = \mathrm{L}_\mathrm{s}^+(f) + \mathrm{L}_\mathrm{s}^+(h) \geq \mathrm{L}_\mathrm{s}^+(f),$$

as required.

QED

In this subsection we do the upgrade from $\mathrm{L}_{\mathrm{s}}^{+} : \mathrm{Bor}_{\mathrm{s}}^{+}(X, \mathbb{R}) \to [0, \infty]$ to $\mathrm{L}^{+} : \mathrm{Bor}^{+}(X, \mathbb{R}) \to [0, \infty]$. The main idea is to *take a* sup.

We will have to adjust this notion a bit, so that we use it in connection to $[0, \infty]$. We wil. thus work with the supremum $\sup(S)$ of a nonempty set $S \subseteq [0, \infty]$, where the equality $\sup(S) = \infty$ means that either $\infty \in S$ or that S is an unbounded subset of $[0, \infty)$.

That being said, we proceed as follows.

Notation 3.5. $\mathrm{L}^{+}(f)$
Let $f \in \mathrm{Bor}^{+}(X, \mathbb{R})$. We define

$$\mathrm{L}^{+}(f) = \sup\{\mathrm{L}^{+}(u) : u \in \mathrm{Bor}_{\mathrm{s}}^{+}(X, \mathbb{R}), u \leq f\}. \tag{3.8}$$

The set appearing in [3.8] is nonempty, as $\underline{0} \leq f$ for any $f \in \mathrm{Bor}^{+}(X, \mathbb{R})$.

Proposition 3.4. L^{+} Extends $\mathrm{L}_{\mathrm{s}}^{+}$
For every $f \in \mathrm{Bor}_{\mathrm{s}}^{+}(X, \mathbb{R})$,

$$\mathrm{L}^{+}(f) = \mathrm{L}_{\mathrm{s}}^{+}(f).$$

Proof. Since $f \leq f$, $\mathrm{L}_{\mathrm{s}}^{+}(f)$ appears in the set in [3.8]. On the other hand, given any $u \in \mathrm{Bor}_{\mathrm{s}}^{+}(X, \mathbb{R})$ with $u \leq f$, $\mathrm{L}_{\mathrm{s}}^{+}(u) \leq \mathrm{L}_{\mathrm{s}}^{+}(f)$ since $\mathrm{L}_{\mathrm{s}}^{+}$ is increasing. This means $\mathrm{L}_{\mathrm{s}}^{+}(f)$ is the supremum of the mentioned set.

QED

We record few basic properties of L^{+}. As we will see, it is quite surprisingly difficult to prove that L^{+} is additive.

Proposition 3.5.
Let $f, g \in \mathrm{Bor}^{+}(X, \mathbb{R})$ and let $\alpha \in [0, \infty]$.

(a) If $f \leq g$, then $\mathrm{L}^{+}(f) \leq \mathrm{L}^{+}(g)$. *increasing*

(b) $\mathrm{L}^{+}(\alpha f) = \alpha\, \mathrm{L}^{+}(f)$. *homogeneity*

Proof Sketch.

(a) This holds since every function from $\{u \in \mathrm{Bor}_{\mathrm{s}}^{+}(X, \mathbb{R}) : u \leq f\}$ is also counted in $\{u \in \mathrm{Bor}_{\mathrm{s}}^{+}(X, \mathbb{R}) : u \leq g\}$.

(b) If $\alpha = 0$, then $\mathrm{L}^{+}(\alpha f) = \mathrm{L}^{+}(\underline{0}) = 0 = 0\, \mathrm{L}^{+}(f)$ (recall $0 \cdot \infty = 0$). If $\alpha > 0$, then it is easy to check that

$$\{v \in \mathrm{Bor}_{\mathrm{s}}^{+}(X, \mathbb{R}) : v \leq \alpha f\} = \{\alpha u : u \in \mathrm{Bor}_{\mathrm{s}}^{+}(X, \mathbb{R}), u \leq f\}.$$

Taking this into the definition of $\mathrm{L}^{+}(\alpha f)$ and using the homogeniety of $\mathrm{L}_{\mathrm{s}}^{+}$, we are done.

QED

The next thing to verify is that L^{+} is additive. Here is half of this verification.

Lemma 3.6.
For every $\mathrm{Bor}^{+}(X, \mathbb{R})$,

$$\mathrm{L}^{+}(f + g) \geq \mathrm{L}^{+}(f) + \mathrm{L}^{+}(g). \tag{3.9}$$

Proof. If $\mathrm{L}^{+}(f + g) = \infty$, then the required inequality is clear; so let us assume $\mathrm{L}^{+}(f + g) < \infty$. This means $\mathrm{L}^{+}(f), \mathrm{L}^{+}(g) < \infty$, since L^{+} is increasing.

So we have assumed that all quantities in [3.9] are finite, and we go to the verification of that inequality via the method of creating an ε of room. It suffices to fix $\varepsilon > 0$ and to prove that

$$\mathrm{L}^+(f+g) \geq \mathrm{L}^+(f) + \mathrm{L}^+(g) - \varepsilon. \tag{3.10}$$

In order to obtain [3.10], we go like this. By the definition of $\mathrm{L}^+(f), \mathrm{L}^+(g)$ as supremums, we can find $u, v \in \mathrm{Bor}^+_{\mathrm{s}}(X, \mathbb{R})$ such that $u \leq f, v \leq g$, and

$$\mathrm{L}^+_{\mathrm{s}}(u) > \mathrm{L}^+(f) - \frac{\varepsilon}{2}, \mathrm{L}^+_{\mathrm{s}}(v) > \mathrm{L}^+(g) - \frac{\varepsilon}{2}.$$

Then $u + v \in \mathrm{Bor}^+_{\mathrm{s}}(X, \mathbb{R})$ is such that $u + v \leq f + g$, with

$$\mathrm{L}^+(f+g) \geq \mathrm{L}^+_{\mathrm{s}}(u+v) = \mathrm{L}^+_{\mathrm{s}}(u) + \mathrm{L}^+_{\mathrm{s}}(v) > \mathrm{L}^+(f) + \mathrm{L}^+(g) - \varepsilon.$$

QED

It may seem like we can do a similar argument with an ε of room in order to prove the inequality opposite to the one stated in Lemma 3.6, saying that

$$\mathrm{L}^+(f+g) \leq \mathrm{L}^+(f) + \mathrm{L}^+(g) \tag{3.11}$$

for all $f, g \in \mathrm{Bor}^+(X, \mathbb{R})$.

But this does not work. The place where things break down is this. Given $w \in \mathrm{Bor}^+_{\mathrm{s}}(X, \mathbb{R})$ such that $w \leq f + g$, it is not clear if and how we could decompose $w = u + v$ with $u, v \in \mathrm{Bor}^+_{\mathrm{s}}(X < \mathbb{R})$ such that $u \leq f, v \leq g$.

However, the formula [3.11] will turn out to hold, in the end – it's only a matter of how we get it. The twist is that we are going to prove the *monotone convergence theorem* (MCT) and only after that we prove the additivity of $\mathrm{L}^+_{\mathrm{s}}$, as an application of the MCT.

Hence we have to leave the inequality [3.11] hanging for a bit, and should first have a look at the MCT. This is a serious tool used in integration theory, which asserts that *one can interchange* L^+ *with an increasing limit*.

In order to make a concise statement of how this exactly goes, let us introduce the following notation.

Notation 3.6. $f_n \nearrow f$

Let $(f_n)_{n=1}^{\infty} \in \left(\mathbb{R}^X\right)^{\mathbb{N}}$ and let $f : X \to \mathbb{R}$. We say $f_n \nearrow f$ to mean that

(a) $f_1 \leq f_2 \leq \cdots$; and *increasing*

(b) for all $x \in X$, $\lim_{n\to\infty} f_n(x) = f(x)$. *pointwise convergence*

Theorem 3.7. Monotone Convergence Theorem

Let $(f_n)_{n=1}^{\infty} \in \mathrm{Bor}^+(X, \mathbb{R})^{\mathbb{N}}$ and let $f \in \mathrm{Bor}^+(X, \mathbb{R})$. If $f_n \nearrow f$, then $\lim_{n\to\infty} \mathrm{L}^+(f_n) = \mathrm{L}^+(f)$.

See Proposition 3.9, 3.10

We conclude this subsection by returning to the discussion of additivity of L^+, by showing how it can be obtained if we assume the MCT. In the next proposition, it is actually nice to take in the full statement about linear combinations.

Proposition 3.8.

For all $f, g \in \mathrm{Bor}^+(X, \mathbb{R}), \alpha, \beta \in [0, \infty)$,

$$\mathrm{L}^+(\alpha f + \beta g) = \alpha\,\mathrm{L}^+(f) + \beta\,\mathrm{L}^+(g).$$

Proof. When we separate the addition and the scalar multiplication that are combined in the statement, and we also take into account the bits of argument provided in Proposition 3.5, 3.6, we see that all we need to prove is

$$\mathrm{L}^+(f+g) \leq \mathrm{L}^+(f) + \mathrm{L}^+(g) \tag{3.12}$$

for all $f, g \in \mathrm{Bor}^+(X, \mathbb{R})$.

In the proof of [3.12] we will assume the MCT, as stated in Theorem 3.7. We will also use the approximation with simple function (see Proposition 2.19).

Let us fix $f, g \in \mathrm{Bor}^+(X,\mathbb{R})$, for which we will prove that [3.12] holds. We know from Proposition 2.19 that we can find sequences $(f_n)_{n=1}^{\infty}, (g_n)_{n=1}^{\infty} \in \mathrm{Bor}_s^+(X,\mathbb{R})^{\mathbb{N}}$ such that $f_n \nearrow f$ and $g_n \nearrow g$. Then $(f_n+g_n)_{n=1}^{\infty}$ is a sequence in $\mathrm{Bor}_s^+(X,\mathbb{R})$ such that $f_n + g_n \nearrow f+g$, and we can write

$$\mathrm{L}^+(f+g) = \lim_{n\to\infty} \mathrm{L}^+(f_n+g_n) = \lim_{n\to\infty} \mathrm{L}_s^+(f_n+g_n) = \lim_{n\to\infty} \mathrm{L}_s^+(f_n) + \mathrm{L}_s^+(g_n) = \lim_{n\to\infty} \mathrm{L}^+(f_n) + \mathrm{L}^+(g_n) \le \mathrm{L}^+(f) + \mathrm{L}^+(g),$$

where the last inequality is from the fact that $f_n \le f, g_n \le g$ and by invoking Proposition 3.5.

QED

3. Proof of Monotone Convergence Theorem

The MCT is an important theorem stated in the previous subsection. It is about the possibility of interchanging the limit with the integral, in a situation where we are dealing with an increasing sequence of functions in $\mathrm{Bor}^+(X,\mathbb{R})$.

We remark that we cannot use the fact that, for all $f, g \in \mathrm{Bor}^+(X,\mathbb{R})$, $\mathrm{L}^+(f+g) \le \mathrm{L}^+(f) + \mathrm{L}^+(g)$ in the proof of MCT, since we proved that result by assuming the MCT.

We first observe an easy statement which makes care of some of the details included in the MCT.

Proposition 3.9.
Let $(f_n)_{n=1}^{\infty} \in \mathrm{Bor}^+(X,\mathbb{R})^{\mathbb{N}}$ and let $f \in \mathrm{Bor}^+(X,\mathbb{R})$. If $f_n \nearrow f$, then

$$\Lambda = \lim_{n\to\infty} \mathrm{L}^+(f_n) \in [0,\infty]$$

exists, with $\Lambda \le \mathrm{L}^+(f)$.

Proof. We have $\underline{0} \le f_1 \le f_2 \le \cdots$, which implies

$$0 \le \mathrm{L}^+(f_1) \le \mathrm{L}^+(f_2) \le \cdots.$$

This means $(\mathrm{L}^+(f_n))_{n=1}^{\infty}$ is an increasing sequence, so has a limit in $[0,\infty]$, as claimed.

Furthermore, from $f_n \nearrow f$, it follows that $f \ge f_n$ for all $n \in \mathbb{N}$. This means $\mathrm{L}^+(f) \ge \mathrm{L}^+(f_n)$ for all $n \in \mathbb{N}$. Thus it follows that

$$\Lambda = \sup_{n\in\mathbb{N}} \mathrm{L}^+(f_n) \le \mathrm{L}^+(f).$$

QED

Having Proposition 3.9 in hand, we now see where is the punch in the MCT, and record that as follows.

Proposition 3.10.
Consider the setting of Proposition 3.9 and the guaranteed limit $\Lambda = \lim_{n\to\infty} \mathrm{L}^+(f_n) \in [0,\infty]$. Then $\Lambda \ge \mathrm{L}^+(f)$.

Postponed

It is clear that when we combine Proposition 3.9, 3.10 we will obtain the MCT. Hence, our job for the rest of this subsection is to prove Proposition 3.10.

An instructive point to consider is this. *What does MCT become if we just look at the indicator functions of sets from $\mathcal{A}$?* Paying attention to indicator function makes sense, since they are the starting point for the construction of L^+.

So suppose that we have $f_n \uparrow f$, where all the f_n's are indicator functions: for all $n \in \mathbb{N}$,

$$f_n = \chi_{A_n}$$

for some $A_n \in \mathcal{A}$. It is immediate that the inequalities $f_1 \le f_2 \le \cdots$ are in this case equivalent to the inclusions $A_1 \subseteq A_2 \subseteq \cdots$. That is, $(A_n)_{n=1}^{\infty} \in \mathcal{A}^{\mathbb{N}}$ is an increasing chain.

Furthermore, the description $f_n \nearrow$ includes the fact that

$$f(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \chi_{A_n}(x)$$

for all $x \in X$. This implies that f it itself an indicator function, namely $f = \chi_{\bigcup_{n=1}^{\infty} A_n}$.

This means the limit $\lim_{n\to\infty} \mathrm{L}^+ (f_n) = \mathrm{L}^+ (f)$ claimed by the MCT amounts in this case to the fact that $\lim_{n\to\infty} \mu (A_n) = \mu (\bigcup_{n=1}^{\infty} A_n)$, which we recognize as the formula for continuity of μ along increasing chains.

Hence we learn is that *MCT is an upgrade of the continuity of μ along increasing chain.* This also gives us a lead on how to prove the MCT: we start from the continuity of μ along increasing chains and build up from there.

We first consider the case when in the setting with $f_n \nearrow f$ of the MCT we insist that f is a simple function – so $f \in \mathrm{Bor}_{\mathrm{s}}^+ (X, \mathbb{R})$ while f_n's can be general functions in $\mathrm{Bor}^+ (X, \mathbb{R})$. This means we will prove the following lemma.

Lemma 3.11.
Let $u \in \mathrm{Bor}_{\mathrm{s}}^+ (X, \mathbb{R})$ and let $(f_n)_{n=1}^{\infty} \in \mathrm{Bor}^+ (X, \mathbb{R})^{\mathbb{N}}$. If $f_n \nearrow u$, then $\lim_{n\to\infty} \mathrm{L}^+ (f_n) \geq \mathrm{L}_{\mathrm{s}}^+ (u)$.

Proof. We assume $u \neq \underline{0}$ for convenience. Indeed, the case $u = \underline{0}$ is trivial.

Since $u \neq \underline{0}$, it assumes some positive values, say $\alpha_1, \ldots, \alpha_p$ (distinct). For every $i \in \{1, \ldots, p\}$, let $A_i = u^{-1} (\alpha_i)$. Then $A_1, \ldots, A_p$ are pairwise disjoint sets in $\mathcal{A}$, with

$$\mathrm{L}_{\mathrm{s}}^+ (u) = \sum_{i=1}^{p} \alpha_i \mu (A_i) . \qquad [3.13]$$

Now we are going to use a trick similar to creating an ε of room. Let us fix $\theta \in (0, 1)$.

In relation to θ, we introduce the following notation. For every $n \in \mathbb{N}$ and every $i \in \{1, \ldots, p\}$, let

$$A_i^{(n)} = f_n^{-1} ([\theta\alpha_i, \infty)) \cap A_i. \qquad [3.14]$$

We have a number of things in relation to these $A_i^{(n)}$'s, which we divide into several claims.

- *Claim 1. Fix $i \in \{1, \ldots, p\}$. Then $\left(A_i^{(n)}\right)_{n=1}^{\infty}$ is an increasing chain of sets from $\mathcal{A}$, whose union is A_i.*

  Proof. By recalling the fact that $(f_n)_{n=1}^{\infty}$ is an increasing sequence, we see that it immediate follows from [3.14] that $\left(A_i^{(n)}\right)_{n=1}^{\infty}$ is an increasing chain. Moreover, $A_i^{(n)} = f_n^{-1} ([\theta\alpha_i, \infty)) \cap A_i$ where $f_n^{-1} ([\theta\alpha_i, \infty)), A_i \in \mathcal{A}$, so that $A_i^{(n)} \in \mathcal{A}$ as well. Lastly, $\bigcup_{n=1}^{\infty} A_i^{(n)} \subseteq A_i$ is clear. On the other hand, given any $x \in A_i$, we know $(f_n (x))_{n=1}^{\infty}$ converges to α_i, meaning that

  $$\lim_{n\to\infty} f_n (x) > \theta\alpha_i.$$

  Thus we can find $n_0 \in \mathbb{N}$ such that $f_{n_0} (x) > \theta\alpha_i$, and for this n_0 we have $x \in A_i^{(n_0)} \subseteq \bigcup_{n=1}^{\infty} A_i^{(n)}$.

- *Claim 2. Let $n \in \mathbb{N}$. Then*

  $$f_n \geq \theta \sum_{i=1}^{p} \alpha_i \chi_{A_i^{(n)}}.$$

  Proof. We have to check that, for all $x \in X$,

  $$f_n (x) \geq \theta \sum_{i=1}^{p} \alpha_i \chi_{A_i^{(n)}} (x) . \qquad [3.15]$$

  To that end, fix $x \in X$. There are two possibilities.

  If $x \notin A_i^{(n)}$ for all $i \in \{1, \ldots, p\}$, then the right-hand side of [3.15] is equal to 0, so we have $f_n (x) \geq 0$, which is clear.

  Suppose otherwise. Then $x \in A_i^{(n)}$ for some unique $i \in \{1, \ldots, p\}$. This is becuse $A_i^{(n)} \subseteq A_i$ for every $i \in \{1, \ldots, p\}$ and $A_1, \ldots, A_p$ are pairwise disjoint. The inequality [3.15] thus amounts to $f_n (x) \geq \theta\alpha_i$, where i is the unique index in $\{1, \ldots, p\}$ such that $x \in A_i^{(n)}$. This is indeed true by definition of $A_i^{(n)}$.

- *Claim 3. Let $n \in \mathbb{N}$. Then*

  $$\mathrm{L}^+ (f_n) \geq \theta \sum_{i=1}^{p} \alpha_i \mu \left(A_i^{(n)}\right).$$

Proof. Note that

$$\mathrm{L}^{+}(f_n) \geq \mathrm{L}^{+}\left(\theta \sum_{i=1}^{p} \alpha_i \chi_{A_i^{(n)}}\right) = \mathrm{L}_{\mathrm{s}}^{+}\left(\theta \sum_{i=1}^{p} \alpha_i \chi_{A_i^{(n)}}\right) = \theta \sum_{i=1}^{p} \alpha_i \, \mathrm{L}_{\mathrm{s}}^{+}\left(\chi_{A_i^{(n)}}\right) = \theta \sum_{i=1}^{p} \alpha_i \mu\left(A_i^{(n)}\right).$$

- *Claim 4.*

$$\lim_{n\to\infty} \mathrm{L}^{+}(f_n) \geq \theta \sum_{i=1}^{p} \alpha_i \mu(A_i).$$

Proof. Take $n \to \infty$ in the inequality from Claim 3.

Upon recalling the formula we had for $\mathrm{L}_{\mathrm{s}}^{+}(u)$ at the beginning of the proof, we see that Claim 4 is giving us the inequality

$$\lim_{n\to\infty} \mathrm{L}^{+}(f_n) \geq \theta \, \mathrm{L}_{\mathrm{s}}^{+}(u). \qquad [3.16]$$

By *unfixing* θ and letting $\theta \to 1$ in [3.16], we reach to the conclusion

$$\lim_{n\to\infty} \mathrm{L}^{+}(f_n) \geq \mathrm{L}_{\mathrm{s}}^{+}(u),$$

which is what we inteded to show.

QED

Now we are ready to prove Proposition 3.10.

Proof of Proposition 3.10

Since $\mathrm{L}^{+}(f)$ is defined as the least upper bound for all the quantities $\mathrm{L}_{\mathrm{s}}^{+}(u)$ where u is a simple Borel function with $u \leq f$, it will suffices to check that Λ is also an upper bound for those quantities.

In short, it suffices to prove that

$$\forall u \in \mathrm{Bor}_{\mathrm{s}}^{+}(X, \mathbb{R}) \left[u \leq f \implies \Lambda \geq \mathrm{L}_{\mathrm{s}}^{+}(u)\right].$$

To that end, fix $u \in \mathrm{Bor}_{\mathrm{s}}^{+}(X, \mathbb{R})$ with $u \leq f$.

From the fact that $f_n \nearrow f$, it follows easily that

$$(f_n \wedge u) \nearrow (f \wedge u).$$

But $f \wedge u = u$, so that

$$(f_n \wedge u) \nearrow u.$$

Lemma 3.11 can be applied in connection to the increasing sequence $(f_n \wedge u)_{n=1}^{\infty}$, since the limit of this sequence is a simple Borel function. From this application of Lemma 3.11, we obtain that

$$\lim_{n\to\infty} \mathrm{L}^{+}(f_n \wedge u) \geq \mathrm{L}_{\mathrm{s}}^{+}(u).$$

But for every $n \geq 1$, we have $f_n \wedge u \leq f_n$, hence $\mathrm{L}_{\mathrm{s}}^{+}(f_n \wedge u) \leq \mathrm{L}^{+}(f_n)$ since L^{+} is increasing. Consequently,

$$\Lambda = \lim_{n\to\infty} \mathrm{L}^{+}(f_n) \geq \lim_{n\to\infty} \mathrm{L}^{+}(f_n \wedge u) \geq \mathrm{L}_{\mathrm{s}}^{+}(u),$$

as desired.

QED

The MCT was stated in a way which forces the increasing sequence $(f_n)_{n=1}^{\infty}$ considered there to have $\lim_{n\to\infty} f_n(x) < \infty$ for all $x \in X$. But suppose we are just given some increasing sequence $(f_n)_{n=1}^{\infty}$ in $\mathrm{Bor}^{+}(X, \mathbb{R})$, where it is allowed that $\lim_{n\to\infty} f_n(x) = \infty$. The MCT stated in Theorem 3.7 does not apply here. However, we can still say something about the increasing limit $\lim_{n\to\infty} \mathrm{L}^{+}(f_n)$. It all depends on whether the set

$$M = \left\{x \in X : \lim_{n\to\infty} f(x) = \infty\right\}$$

has a measure zero or not. This will be addressed in a more precise way in an assignment.

Fix a measure space $(X, \mathcal{A}, \mu)$. Recall that we have $\operatorname{Bor}(X, \mathbb{R})$, the unital algebra and lattice of $\mathcal{A}/\mathcal{B}_{\mathbb{R}}$-measurable functions. We also have inclusions

$$\operatorname{Bor}(X, \mathbb{R}) \supseteq \operatorname{Bor}^+(X, \mathbb{R}) \supseteq \operatorname{Bor}_s^+(X, \mathbb{R}),$$

where we started our construction of integration from $\operatorname{Bor}_s^+(X, \mathbb{R})$ by defining $\mathrm{L}_s^+ : \operatorname{Bor}_s^+(X, \mathbb{R}) \to [0, \infty]$. We furthermore take the supremum of suitable quantities concerning L_s^+ to define $\mathrm{L}^+ : \operatorname{Bor}^+(X, \mathbb{R}) \to [0, \infty]$, which extends L_s^+. What we now want to define is L that extends L^+ to $\operatorname{Bor}(X, \mathbb{R})$. However, it turns out if we are to define such L to every Borel functions, then we face the *forbidden operation* $\infty - \infty$. Hence we are going to define a subspace $\mathcal{L}^1(\mu)$ of $\operatorname{Bor}(X, \mathbb{R})$ on which we can safely define L.

Notation 3.7. f_+, f_-

Let $f : X \to \mathbb{R}$. We are going to write

$$\begin{aligned} f_+ &= f \vee \underline{0} \\ f_- &= (-f) \vee \underline{0} \end{aligned}.$$

Observe that

$$\begin{aligned} f_+ + f_- &= f \\ f_+ - f_- &= |f| \end{aligned}.$$

It is immediate from the definition that $f_+, f_- \in \operatorname{Bor}^+(X, \mathbb{R})$ whenever $f \in \operatorname{Bor}(X, \mathbb{R})$.

Lemma 3.12.

Let $f \in \operatorname{Bor}(X, \mathbb{R})$. The following are equivalent.

(a) $\mathrm{L}^+(|f|) < \infty$.

(b) $\mathrm{L}^+(f_+) < \infty$ and $\mathrm{L}^+(f_-) < \infty$.

Proof. If $\mathrm{L}^+(|f|) < \infty$, then we have $\mathrm{L}^+(f_+) \le \mathrm{L}^+(|f|) < \infty$ and similarly $\mathrm{L}^+(f_-) \le \mathrm{L}^+(|f|) < \infty$.

Conversely, if $\mathrm{L}^+(f_+), \mathrm{L}^+(f_-) < \infty$, then $\mathrm{L}^+(|f|) = \mathrm{L}^+(f_+) - \mathrm{L}^+(f_-) < \infty$.

QED

We are now ready to define what $\mathcal{L}^1(\mu)$ is.

Def'n 3.8. **Integrable** Function

Let

$$\mathcal{L}^1(\mu) = \{f \in \operatorname{Bor}(X, \mathbb{R}) : \mathrm{L}^+(|f|) < \infty\}.$$

We call each function in $\mathcal{L}^1(\mu)$ as an ***integrable*** function.

Notation 3.9. $\mathrm{L}(f)$

For every $f \in \mathcal{L}^1(\mu)$, we denote

$$\mathrm{L}(f) = \mathrm{L}^+(f_+) - \mathrm{L}^+(f_-) \in \mathbb{R}.$$

So far we have two functionals $\mathrm{L} : \mathcal{L}^1(\mu) \to \mathbb{R}, \mathrm{L}^+ : \operatorname{Bor}^+(X, \mathbb{R}) \to [0, \infty]$ where both are trying to be the *Lebesgue integral.* Observe that, for any $f \in \mathcal{L}^1(\mu) \cap \operatorname{Bor}^+(X, \mathbb{R})$, we have

$$\mathrm{L}(f) = \mathrm{L}^+(f),$$

since $f \ge \underline{0}$ implies $f_+ = f, f_- = \underline{0}$, so that

$$\mathrm{L}(f) = \mathrm{L}^+(f_+) - \mathrm{L}^+(f_-) = \mathrm{L}^+(f) - \mathrm{L}^+(\underline{0}) = \mathrm{L}^+(f).$$

Hence L, L^+ both agree for such f.

Lemma 3.13.
Let $f \in \mathcal{L}^1(\mu)$ be such that $f = h_1 - h_2$ for some $h_1, h_2 \in \mathrm{Bor}^+(X, \mathbb{R})$ with $\mathrm{L}^+(h_1), \mathrm{L}^+(h_2) < \infty$. Then

$$\mathrm{L}(f) = \mathrm{L}^+(h_1) - \mathrm{L}^+(h_2).$$

Proof. Note that

$$f_+ - f_- = f = h_1 - h_2,$$

rearranging which gives

$$h_1 + f_- = h_2 + f_+.$$

Hence

$$\mathrm{L}^+(h_1) + \mathrm{L}^+(f_-) = \mathrm{L}^+(h_1 + f_-) = \mathrm{L}^+(h_2 + f_+) = \mathrm{L}^+(h_2) + \mathrm{L}^+(f_+),$$

so that

$$\mathrm{L}(f) = \mathrm{L}^+(f_+) - \mathrm{L}^+(f_-) = \mathrm{L}^+(h_1) - \mathrm{L}^+(h_2).$$

QED

Proposition 3.14.

(a) $\mathcal{L}^1(\mu)$ is a vector subspace of $\mathrm{Bor}(X, \mathbb{R})$.
(b) $\mathrm{L} : \mathcal{L}^1(\mu) \to \mathbb{R}$ is linear.

Proof of (a). We need to check that that $\mathcal{L}^1(\mu)$ is closed under linear combination and has $\underline{0}$. Clearly $\underline{0} \in \mathcal{L}^1(\mu)$. We note that given any $f, g \in \mathcal{L}^1(\mu)$ and $\alpha, \beta \in \mathbb{R}$,

$$\mathrm{L}^+(|\alpha f + \beta g|) < \infty$$

by observing that

$$|\alpha f(x) + \beta g(x)| \le |\alpha f(x)| + |\beta g(x)| = |\alpha| |f(x)| + |\beta| |g(x)|$$

for all $x \in \mathbb{R}$, so that

$$\mathrm{L}^+(|\alpha f + \beta g|) \le \mathrm{L}^+(|\alpha| |f| + |\beta| |g|) = \mathrm{L}^+(|\alpha| |f|) + \mathrm{L}^+(|\beta| |g|) = |\alpha| \mathrm{L}^+(|f|) + |\beta| \mathrm{L}^+(|g|) < \infty.$$

Proof of (b). We split the proof into two parts. Let $f, g \in \mathcal{L}^1(\mu), \alpha \in \mathbb{R}$.

- *Claim 1.* $\mathrm{L}(f+g) = \mathrm{L}(f) + \mathrm{L}(g)$.

  Proof. Write $f = f_+ - f_-, g = g_+ - g_-$. Then we have

$$\mathrm{L}(f) = \mathrm{L}^+(f_+) - \mathrm{L}^+(f_-),$$
$$\mathrm{L}(g) = \mathrm{L}^+(g_+) - \mathrm{L}^+(g_-).$$

  Now write

$$f + g = (f_+ - f_-) + (g_+ - g_-) = \underbrace{(f_+ + g_+)}_{=h_1} - \underbrace{(f_- + g_-)}_{=h_2} = h_1 - h_2,$$

  where $h_1, h_2 \in \mathrm{Bor}^+(X, \mathbb{R})$, with $\mathrm{L}^+(h_1) = \mathrm{L}^+(f_+) + \mathrm{L}^+(g_+) < \infty$ and $\mathrm{L}^+(h_2) = \mathrm{L}^+(f_-) + \mathrm{L}^+(g_-) < \infty$. Hence

$$\mathrm{L}(f+g) = \mathrm{L}^+(h_1) - \mathrm{L}^+(h_2) = \mathrm{L}^+(f_+) + \mathrm{L}^+(g_+) - \mathrm{L}^+(f_-) + \mathrm{L}^+(g_-) = \mathrm{L}(f) + \mathrm{L}(g).$$

- *Claim 2.* $\mathrm{L}(\alpha f) = \alpha \mathrm{L}(f)$.

  Proof. When $\alpha = 0$, the result is clear. Suppose $\alpha > 0$. Note that

$$\mathrm{L}(\alpha f) = \mathrm{L}^+(\alpha f_+) - \mathrm{L}^+(\alpha f_-) = \alpha \mathrm{L}^+(f_+) - \alpha \mathrm{L}^+(f_-) = \alpha(\mathrm{L}^+(f_+) - \mathrm{L}^+(f_-)) = \alpha \mathrm{L}(f).$$

  When $\alpha < 0$, we note that

$$\mathrm{L}(\alpha f) = \mathrm{L}^+(-\alpha f_-) - \mathrm{L}^+(-\alpha f_+) = -\alpha(\mathrm{L}^+(f_-) - \mathrm{L}^+(f_+)) = -\alpha(-\mathrm{L}(f)) = \mathrm{L}(f).$$

QED

Proposition 3.15.

(a) If $f, g \in \mathcal{L}^1(\mu)$ are such that $f \leq g$, then $\mathrm{L}(f) \leq \mathrm{L}(g)$.

(b) If $f \in \mathcal{L}^1(\mu)$, then $|f| \in \mathcal{L}^1(\mu)$ with

$$|\mathrm{L}(f)| \leq \mathrm{L}(|f|).$$

Proof of (a). Observe that

$$\mathrm{L}(g) - \mathrm{L}(f) = \mathrm{L}(g - f) = \mathrm{L}^+(g - f) \geq 0.$$

Proof of (b). It suffices to note that $-|f| \leq f \leq |f|$.

QED

Notation 3.10. $\int f \, \mathrm{d}\mu$.

Let $f \in \mathrm{Bor}(X, \mathbb{R})$, for which at least one of $\mathrm{L}^+(f), \mathrm{L}(f)$. Then we write

$$\int f \, \mathrm{d}\mu = \int_X f(x) \, \mathrm{d}\mu(x)$$

for $\mathrm{L}^+(f)$ or $\mathrm{L}(f)$.

We remark that the above quantity $\int_X f \, \mathrm{d}\mu$ is well-defined even if both $\mathrm{L}^+(f), \mathrm{L}(f)$ are defined, since we know

$$\mathrm{L}^+(f) = \mathrm{L}(f).$$

Suppose $A \in \mathcal{A}$ is nonempty. Then *what does it mean by writing $\int_A f(x) \, \mathrm{d}\mu(x)$?* There are two possible answers.

(a) $\int_A f(x) \, \mathrm{d}\mu(x) = \int_X f(x) \chi_A(x) \, \mathrm{d}\mu(x)$.

(b) We can consider the restricted measure space $(A, \mathcal{A}_0, \mu_0)$ and the restriction $f_0 = f|_A$, and write

$$\int_A f(x) \, \mathrm{d}\mu(x) = \int_A f_0(x) \, \mathrm{d}\mu_0(x).$$

As we will see in an assignment, these two methods *coincide*.

5. Lebesgue's Dominated Convergence Theorem

We again consider the space of functions $\mathcal{L}^1(\mu)$ introduced in the previous subsection, and we will prove an important theorem, the *Lebesgue's dominated convergence theorem* (LDCT), which addresses the issue of when we can *interchange integration with a limit*.

Interchanging integration with a limit really has issues, and it is easy to point out examples where the two quantities are not equal to each other. The insight of Lebesgue was to spot a sufficient condition which clears these issues – and where, moreover, the condition is easy to verify and also turns out to hold in many applications. In words, this sufficient condition says that, the functions f_n appearing in a sequence $(f_n)_{n=1}^\infty \in \mathrm{Bor}(X, \mathbb{R})^{\mathbb{N}}$ admit a *dominating function* which is integrable.

This is made formal in below.

Theorem 3.16. Lebesgue's Dominated Convergence Theorem (LDCT)

Let $(f_n)_{n=1}^{\infty} \in \mathrm{Bor}(X, \mathbb{R})$ such that $f_n \to f \in \mathrm{Bor}(X, \mathbb{R})$ pointwise and suppose there exists $h \in \mathcal{L}^1(\mu) \cap \mathrm{Bor}^+(X, \mathbb{R})$ which dominates each f_n, in the sense that $|f_n| \leq h$. Then $f, f_1, \ldots \in \mathcal{L}^1(\mu)$ and

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

The LDCT can actaully be stated in a form which looks a bit different, in connection to a map $\|\cdot\|_1 : \mathcal{L}^1(\mu) \to [0, \infty)$. After defining $\|\cdot\|_1$, we will rephrase LDCT in a way which refers to $\|\cdot\|_1$.

Notation 3.11. $\|f\|_1$

For every $f \in \mathcal{L}^1(\mu)$, we define

$$\|f\|_1 = \int |f| \, \mathrm{d}\mu \in [0, \infty). \tag{3.17}$$

The integral on the right-hand side of [3.17] can be viewed either in the L^+ sense of in the L sense, since $|f| \in \mathcal{L}^1(\mu) \cap \mathrm{Bor}^+(X, \mathbb{R})$.

Note that $\|\cdot\|_1$ already appeared in the previous subsection (see Proposition 3.15). The inequality in (b) of Proposition 3.15 can be presented in the form

$$\left| \int f \, \mathrm{d}\mu \right| \leq \|f\|_1 \tag{3.18}$$

for all $f \in \mathcal{L}^1(\mu)$.

It is easily verified that $\|\cdot\|_1$ is a semi-norm on $\mathcal{L}^1(\mu)$, which means

$$\|f + g\|_1 \leq \|f\|_1 + \|g\|_1$$

and

$$\|af\|_1 = |a| \, \|f\|_1$$

for all $f, g \in \mathcal{L}^1(\mu), a \in \mathbb{R}$.

Now back to the LDCT.

Theorem 3.17. Restatement of LDCT with respect to $\|\cdot\|_1$

Let $(f_n)_{n=1}^{\infty} \in \mathrm{Bor}(X, \mathbb{R})^{\mathbb{N}}$ such that $\lim_{n\to\infty} f_n = f$ pointwise for some $f \in \mathrm{Bor}(X, \mathbb{R})$. If there exists $h \in \mathcal{L}^1(\mu) \cap \mathrm{Bor}^+(X, \mathbb{R})$ such that $|f_n| \leq h$ for all $n \in \mathbb{N}$, then $f, f_1, f_2, \ldots \in \mathcal{L}^1(\mu)$, with

$$\lim_{n\to\infty} \|f_n - f\|_1 = 0.$$

Theorem 3.16, 3.17 have the same hypothesis, and it is easy to see that the conclusion of Theorem 3.17 implies that of Theorem 3.16.

Lemma 3.18.

Suppose $(f_n)_{n=1}^{\infty} \in \mathcal{L}^1(\mu)^{\mathbb{N}}$ is such that

$$\lim_{n\to\infty} \|f - f_n\|_1 = 0 \tag{3.19}$$

for some $f \in \mathcal{L}^1(\mu)$. Then

$$\lim_{n\to\infty} \int f_n \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu. \tag{3.20}$$

Proof. Note that

$$0 = \lim_{n\to\infty} \|f - f_n\|_1 \geq \lim_{n\to\infty} \left| \int f - f_n \, \mathrm{d}\mu \right| = \lim_{n\to\infty} \left| \int f \, \mathrm{d}\mu - \int f_n \, \mathrm{d}\mu \right| \qquad [3.21]$$

by linearity of the Lebesgue integral and [3.18]. But [3.21] means $\lim_{n\to\infty} \left|\int f_n \, \mathrm{d}\mu - \int f \, \mathrm{d}\mu\right| = 0$, which precisely means [3.20].

QED

Notation 3.12. $f_n \searrow f$

Let $(f_n)_{n=1}^{\infty} \in \left(\mathbb{R}^X\right)^{\mathbb{N}}$ and let $f : X \to \mathbb{R}$. We say $f_n \searrow f$ to mean that

(a) $f_1 \geq f_2 \geq \cdots$; and *decreasing*

(b) for all $x \in X$, $\lim_{n\to\infty} f_n(x) = f(x)$. *pointwise convergence*

Proposition 3.19.

Let $(u_n)_{n=1}^{\infty} \in \mathrm{Bor}^+(X, \mathbb{R})^{\mathbb{N}}$ be such that $u_n \searrow u$ for some $u \in \mathrm{Bor}^+(X, \mathbb{R})$. If $\int u_1 \, \mathrm{d}\mu < \infty$, then

$$\lim_{n\to\infty} \int u_n \, \mathrm{d}\mu = \int u \, \mathrm{d}\mu.$$

Proof. For every $n \in \mathbb{N}$, let

$$f_n = u_1 - u_n.$$

Since $(u_n)_{n=1}^{\infty}$ is a decreasing sequence in $\mathrm{Bor}^+(X, \mathbb{R})$, $(f_n)_{n=1}^{\infty}$ is an increasing sequence in $\mathrm{Bor}^+(X, \mathbb{R})$. Since $u_n \to u$ pointwise, $f_n = u_1 - u_n \to u_1 - u$ pointwise. Hence by the MCT,

$$\lim_{n\to\infty} \int f_n \, \mathrm{d}\mu = \int u_1 - u \, \mathrm{d}\mu,$$

rearranging which gives

$$\lim_{n\to\infty} \int u_1 - f_n = \int u \, \mathrm{d}\mu.$$

Thus

$$\lim_{n\to\infty} \int u_n \, \mathrm{d}\mu = \lim_{n\to\infty} \int u_1 - f_n \, \mathrm{d}\mu = \int u \, \mathrm{d}\mu,$$

as required.

QED

Proposition 3.20.

Let $(g_n)_{n=1}^{\infty} \in \mathrm{Bor}^+(X, \mathbb{R})$ be such that $g_n \to \underline{0}$ pointwise. Suppose moreover that there exists $h \in \mathrm{Bor}^+(X, \mathbb{R}) \cap \mathcal{L}^1(\mu)$ such that $g_n \leq h$ for all $n \in \mathbb{N}$. Then

$$\lim_{n\to\infty} \int g_n \, \mathrm{d}\mu = 0.$$

Proof. For each $n \in \mathbb{N}$, let $u_n : X \to \mathbb{R}$ be defined by

$$u_n(x) = \sup_{k \geq n} g_k(x) \leq h(x) \in [0, \infty). \qquad [3.22]$$

We claim that $u_n \in \mathrm{Bor}^+(X, \mathbb{R})$ for all $N \in \mathbb{N}$. To verify this, suppose $a \in \mathbb{R}$ is given, for which we verify $u_n^{-1}((a, \infty)) \in \mathcal{A}$; this will prove that u_n is $\mathcal{A} / \mathcal{B}_{\mathbb{R}}$-measurable. Note that

$$\begin{aligned} u_n^{-1}((a, \infty)) &= \{x \in X : u_n(x) \in (a, \infty)\} \\ &= \{x \in X : \exists k \in \mathbb{N}\, [g_k(x) \geq a]\} \\ &= \bigcup_{k \geq n} g_k^{-1}((a, \infty)). \end{aligned}$$

Since $g_k \in \mathrm{Bor}^+(X, \mathbb{R})$, it follows that $g_k^{-1}((a, \infty)) \in \mathcal{A}$ for all $k \geq n$. Since $\mathcal{A}$ is closed under countable unions, we conclude $u^{-1}((a, \infty)) \in \mathcal{A}$. Since $u \geq \underline{0}$ is clear from the definition, $u \in \mathrm{Bor}^+(X, \mathbb{R})$, as we claimed (in fact, see Proposition 2.18).

Now, $(u_n)_{n=1}^{\infty}$ is a decreasing sequence by definition. Moreover,

$$\inf_{n \in \mathbb{N}} u_n(x) = \inf_{n \in \mathbb{N}} \sup_{k \geq n} g_k(x) = \limsup_{n \to \infty} g_n(x) = \lim_{n \to \infty} g_n(x) = 0$$

for all $x \in X$, since $g_n \to \underline{0}$ pointwise. This means $u_n \searrow \underline{0}$. Moreover, it is clear from [3.22] that $\int u_1 \, \mathrm{d}\mu \leq \int h \, \mathrm{d}\mu < \infty$, since $h \in \mathcal{L}^1(\mu)$. Therefore, by invoking Proposition 3.19 with respect to $(u_n)_{n=1}^{\infty}$, we obtain

$$\lim_{n \to \infty} \int u_n \, \mathrm{d}\mu = \int \underline{0} \, \mathrm{d}\mu = 0.$$

But then

$$\lim_{n \to \infty} \int g_n \, \mathrm{d}\mu \leq \lim_{n \to \infty} \int_X \sup_{k \geq n} g_k(x) \, \mathrm{d}\mu = \lim_{n \to \infty} \int u_n \, \mathrm{d}\mu = 0,$$

so that

$$\lim_{n \to \infty} \int g_n \, \mathrm{d}\mu = 0,$$

as required.

QED

Example 3.3.

Consider the setting $(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \lambda)$ and for all $n \in \mathbb{N}$, let $f_n = \chi_{(n,n+1)}$. Then note that $f_n \in \mathrm{Bor}_s^+(\mathbb{R}, \mathbb{R})$ for all $n \in \mathbb{N}$,

$$\int f_n \, \mathrm{d}\lambda = 1,$$

and $\lim_{n \to \infty} f_n(t) = 0$ for all $t \in \mathbb{R}$. This means

$$\lim_{n \to \infty} \int f_n \, \mathrm{d}\lambda = 1 \neq 0 \int \lim_{n \to \infty} f_n \, \mathrm{d}\lambda.$$

Does this contradict LDCT?

Answer. No; note that, in order for a fuction $h : \mathbb{R} \to \mathbb{R}$ to dominate all f_n's, $h(x) \geq 1$ for all $x > 1$. This means $h \notin \mathcal{L}^1(\lambda)$.

QED

6. Almost Everywhere

We again fix a measure space $(X, \mathcal{A}, \mu)$.

Def'n 3.13. **Almost Everywhere** Predicate

Consider a predicate P over X. We say P is (true) ***almost everywhere*** with respect to μ, often abbreviated as P is ***a.e.***-μ, to mean that there exists a negligible set $N \in \mathcal{A}$ such that $P(x)$ is true for all $x \in X \setminus N$.

Proposition 3.21. A Characterization of a.e.-μ Equal Functions

Let $f, g \in \mathrm{Bor}(X, \mathbb{R})$ and let $M = \{x \in X : f(x) \neq g(x)\}$.

(a) $M \in \mathcal{A}$.

(b) $f = g$ a.e.-μ if and only if $\mu(M) = 0$.

Proof.

(a) Write

$$M = \{x \in X : f(x) \neq g(x)\} = \{x \in X : (f - g)(x) \neq 0\} = (f - g)^{-1}(\mathbb{R} \setminus \{0\}).$$

Since $f - g \in \mathrm{Bor}(X, \mathbb{R})$ and $\mathbb{R} \setminus \{0\} \in \mathcal{B}_{\mathbb{R}}$, it follows that $A \in \mathcal{A}$.

(b) Suppose $f = g$ a.e.-μ. This means there is $N \in \mathcal{A}$ such that $\mu(N) = 0$ and $f(x) = g(x)$ for all $x \in X \setminus N$. But for any $x \in X$,

$$x \in M \implies f(x) \neq g(x) \implies x \notin X \setminus N \implies x \in N.$$

Hence $M \subseteq N$, so that $0 \leq \mu(M) \leq \mu(N) = 0$, forcing $\mu(M) = 0$.

For the converse direction, we note that by picking $N = M$ the conditions in Def'n 3.13 are satisfied.

QED

We have alternative way of characterizing $f = g$ a.e.-μ.

Lemma 3.22.
Let $f \in \mathrm{Bor}^+(X, \mathbb{R})$ and consider $\int_X f \, \mathrm{d}\mu \in [0, \infty]$. The following are equivalent.

(a) $f = 0$ a.e.-μ.

(b) $\int_X f \, \mathrm{d}\mu = 0$.

Proof. We are going to use the definition of L^+ directly. The following claim is what makes our argument work.

We claim that, *for every $u \in \mathrm{Bor}_s^+(X, \mathbb{R})$ such that $u \leq f$, we have $\int_X u \, \mathrm{d}\mu = 0$.* To verify this, we observe that

$$u^{-1}(0) \supseteq f^{-1}(0),$$

since both u, f are nonnegative and $u \leq f$. So let $Y = u^{-1}(0)$, which is a set in $\mathcal{A}$ since $u \in \mathrm{Bor}(X, \mathbb{R})$. This means $X \setminus Y \in \mathcal{A}$ as well, with

$$X \setminus Y = X \setminus u^{-1}(0) \subseteq X \setminus f^{-1}(0).$$

But $f = 0$ a.e.-μ, which precisely means $\mu\left(X \setminus f^{-1}(0)\right) = 0$. Consequently, $\mu(X \setminus Y) = 0$. Now, given any $a > 0$ that u assumes, $u^{-1}(a) \subseteq X \setminus u^{-1}(0) = X \setminus Y$, so that $\mu\left(u^{-1}(a)\right) = 0$. Thus it follows from the definition of the Lebesgue integral for simple nonnegative functions that $\int_X u \, \mathrm{d}\mu = 0$, as claimed.

By recalling that $\mathrm{L}^+(f)$ is simply the supremum of the simple functions which *underestimates* f, we conclude $\mathrm{L}^+(f) = 0$. This ends (a) $\implies$ (b).

For the converse direction, we claim that, *for all $\varepsilon > 0$, the set*

$$S_\varepsilon = \{x \in X : f(x) \geq \varepsilon\}$$

has measure zero, with respect to μ. This is rather straightforward, since if we assume $\mu(S_\varepsilon) > 0$ for some $\varepsilon > 0$, then by defining

$$u = \begin{cases} 0 & \text{if } x \notin S_\varepsilon \\ \varepsilon & \text{if } x \in S_\varepsilon \end{cases}$$

for all $x \in X$, $u \leq f$ with $\int_X u \, \mathrm{d}\mu = \mu(S_\varepsilon)\,\varepsilon \neq 0$. This contradicts $\int_X f \, \mathrm{d}\mu = 0$.

Now we consider the particular case where $\varepsilon \in \left\{1, \frac{1}{2}, \ldots\right\}$, so that we can write

$$S = \bigcup_{n \in \mathbb{N}} S_{\frac{1}{n}}$$

and ensure that $S \in \mathcal{A}$, as a countable union of measurable sets. But note that, by defnintion, for any $x \in S$, we have that $f(x) \geq \frac{1}{n}$ for some $n \in \mathbb{N}$, so that $f(x) > 0$. Conversely, if $x \in X$ is such that $f(x) > 0$, then $f(x) > \frac{1}{n}$ for some $n \in \mathbb{N}$. This means

$$S = \{x \in X : f(x) > 0\}.$$

But we noted that $\mu(S_\varepsilon) = 0$ for all $\varepsilon > 0$, so in particular $\mu\left(S_{\frac{1}{n}}\right) = 0$ for all $n \in \mathbb{N}$. We also see from the definition that $S_{\frac{1}{n}} \subseteq S_{\frac{1}{n+1}}$ for all $n \in \mathbb{N}$. This means $\left(S_{\frac{1}{n}}\right)_{n=1}^{\infty}$ is an increasing chain of sets in $\mathcal{A}$, a perfect setting to apply the *continuity of μ along increasing chains*, obtaining

$$\mu(S) = \lim_{n \to \infty} \mu\left(S_{\frac{1}{n}}\right) = \lim_{n \to \infty} 0 = 0.$$

This S is by construction $X \setminus f^{-1}(0)$, so we thus conclude $f = 0$ a.e.-μ.

QED

Proposition 3.23.

Let $f, g \in \mathrm{Bor}(X, \mathbb{R})$ and consider the function $|f - g| \in \mathrm{Bor}^+(X, \mathbb{R})$. The following are equivalent.

(a) $f = g$ a.e.-μ.

(b) $\int_X |f - g| \, \mathrm{d}\mu = 0$.

Proof. Denoting $h = |f - g| \in \mathrm{Bor}^+(X, \mathbb{R})$, we see that

$$f = g \text{ a.e.-}\mu \iff h = 0 \text{ a.e.-}\mu.$$

Upon applying Lemma 3.22 to h we thus find that

$$f = g \text{ a.e.-}\mu \iff h = 0 \text{ a.e.-}\mu \iff \int_X h \, \mathrm{d}\mu = 0 \iff \int_X |f - g| \, \mathrm{d}\mu = 0,$$

as required.

QED

Now let us look at the case when we consider functions from the space $\mathcal{L}^1(\mu)$. Recall that we define $\|\cdot\|_1 : \mathcal{L}^1(\mu) \to [0, \infty)$ by

$$\|f\|_1 = \int_X f \, \mathrm{d}\mu$$

for all $f \in \mathcal{L}^1(\mu)$. From Proposition 3.23, we thus infer the following.

Corollary 3.23.1.

Let $f, g \in \mathcal{L}^1(\mu)$. The following are equivalent.

(a) $f = g$ a.e.-μ.

(b) $\|f - g\|_1 = 0$.

See Proposition 3.24

In fact in the setting of Corollary 3.23.1, it actually suffices if we assume one of the two functions f, g to be in $\mathcal{L}^1(\mu)$. This is recorded in the next proposition.

Proposition 3.24.

Let $f, g \in \mathrm{Bor}(X, \mathbb{R})$ such that $f = g$ a.e.-μ. If one of f, g is in $\mathcal{L}^1(\mu)$, then so is the other. Moreover, we have

$$\int_X f \, \mathrm{d}\mu = \int_X g \, \mathrm{d}\mu$$

and

$$\|f\|_1 = \|g\|_1.$$

Proof. Since $f, g \in \mathrm{Bor}(X, \mathbb{R})$ are such that $f = g$ a.e.-μ, Proposition 3.23 can be applied (without assuming that f or g are integrable) to infer that $f - g \in \mathcal{L}^1(\mu)$ and $\|f - g\|_1 = 0$.

We are also told that one of f, g is integrable; without loss of generality assume $f \in \mathcal{L}^1(\mu)$. Then we get the integrability of g by writing $g = f - (f - g)$ with $f, f - g \in \mathcal{L}^1(\mu)$ and by invoking the closure of $\mathcal{L}^1(\mu)$ under linear combinations.

By starting from the fact that $\|f - g\|_1 = 0$ we write that

$$0 \leq \left| \int_X f \, \mathrm{d}\mu - \int_X g \, \mathrm{d}\mu \right| = \left| \int_X (f - g) \, \mathrm{d}\mu \right| \leq \|f - g\|_1 = 0,$$

which implies that $\int_X f \, \mathrm{d}\mu = \int_X g \, \mathrm{d}\mu$.

Moreover,

$$\|f\|_1 = \|(f-g)+g\|_1 \leq \|f-g\|_1 + \|g\|_1 = 0 + \|g\|_1$$

to obtain that $\|f\|_1 \leq \|g\|_1$. A similar argument shows that $\|g\|_1 \leq \|f\|_1$, and we conclude htat $\|f\|_1 = \|g\|_1$, as required.

QED

It is easy to verify that the notion of equality a.e.-μ gives an equivalence relatoin on $\mathrm{Bor}(X,\mathbb{R})$ – that is, it enjoys the three properties of being reflexive, symmetric and transitive.

An equivalence relation breaks the underlying set into equivalence classes. In the case at hand, $\mathrm{Bor}(X,\mathbb{R})$ is therefore decomposed into a disjoint union of equivalence classes, where two functions $f, g \in \mathrm{Bor}(X,\mathbb{R})$ belong to the same class if and only if $f = g$ a.e.-μ. In the language of these equivalence classes, Proposition 3.24 can be stated as saying that

$\mathcal{L}^1(\mu)$ is *saturated* with respect to equality a.e.-μ.

That is, whenever $f \in \mathcal{L}^1(\mu)$, the whole equivalence class of f is contained in $\mathcal{L}^1(\mu)$.

Many textbooks go for *identifying functions* which are equal a.e.-μ. Formally, what they do goes under the name of *taking the quotient* of $\mathrm{Bor}(X,\mathbb{R})$ by the equivalence relation given by equality a.e.-μ. Doing such identifications has some merits, but also generates the issue that our basic notion of *what is a function* starts to depend on what measure μ we are consindering on a measurable space $(X,\mathcal{A})$. We will generally avoid this – we will rather prefer to stick to the *usual* notion of functoin.

7. Lebesgue Integration and Riemann Integration on a Finite Interval

We fix $a < b$ in $\mathbb{R}$ and we look at the measure space $\left([a,b], \mathcal{B}_{[a,b]}, \lambda_{[a,b]}\right)$. In connection to this measure space, we are going to introduce the following notation.

Notation 3.14. $\mathrm{Bor}_b([a,b],\mathbb{R})$
We will denote

$$\mathrm{Bor}_b([a,b],\mathbb{R}) = \{f \in \mathrm{Bor}([a,b],\mathbb{R}) : f \text{ is bounded}\}.$$

It is immediate that $\mathrm{Bor}_b([a,b],\mathbb{R})$ is a linear subspace of $\mathrm{Bor}([a,b],\mathbb{R})$ which is also closed under multiplication, under $\vee$ and $\wedge$ operations, and under the operation of taking the absolute value.

An important point to notice is that we have $\mathrm{Bor}_b([a,b],\mathbb{R}) \subseteq \mathcal{L}^1\left(\lambda_{[a,b]}\right)$. Indeed, if $f \in \mathrm{Bor}_b([a,b],\mathbb{R})$, then we can pick an $r > 0$ such that $|f(x)| \leq r$ for all $x \in [a,b]$, and it follows that

$$\int_{[a,b]} |f| \, \mathrm{d}\lambda_{[a,b]} \leq \int_{[a,b]} r \, \mathrm{d}\lambda_{[a,b]} = r(b-a) < \infty,$$

showing that $f \in \mathcal{L}^1\left(\lambda_{[a,b]}\right)$. Thus we are sure we can talk here about Lebesgue integrals.

Fix $f \in \mathrm{Bor}_b([a,b],\mathbb{R})$ and let us do a quick review from calculus, concerning the fact that f has an upper Riemann integral and a lower Riemann integral over $[a,b]$. This goes as follows.

(a) We consider *divisions* of the interval $[a,b]$, which are systems of points

$$\Delta = (t_0, \ldots, t_k)$$

with $k \geq 1$ and $a = t_0 < t_1 < \cdots < t_k = b$.

(b) Given a division $\Delta = (t_0, t_1, \ldots, t_k)$ of $[a,b]$, we define the *upper Darboux sum* $U(f,\Delta)$ and *lower Darboux sum* $L(f,\Delta)$ associated to f and Δ, which are

$$U(f,\Delta) = \sum_{i=1}^{k} (t_i - t_{i-1}) \sup_{t_{i-1} \leq t \leq t_i} f(t)$$

and

$$L(f,\Delta) = \sum_{i=1}^{k} (t_i - t_{i-1}) \inf_{t_{i-1} \leq t \leq t_i} f(t).$$

(c) We then define the *upper integral* $\overline{\int_a^b} f(t)\ \mathrm{d}t$ and *lower integral* $\underline{\int_a^b} f(t)\ \mathrm{d}t$ of f on $[a,b]$ to be

$$\overline{\int_a^b} f\,\mathrm{d}t = \inf\{U(f,\Delta) : \Delta \text{ is a division of } [a,b]\}$$

and

$$\underline{\int_a^b} f\,\mathrm{d}t = \sup\{U(f,\Delta) : \Delta \text{ is a division of } [a,b]\}.$$

We know from calculus that the upper and lower integrals are well-defined and we always have the inequality

$$\underline{\int_a^b} f(t)\ \mathrm{d}t \leq \overline{\int_a^b} f(t)\ \mathrm{d}t. \qquad [3.23]$$

If [3.23] holds with an equality, then we declare f to be *Riemann integrable*, and we define the Riemann integral of f to be the common value of the pper and lower integrals:

$$(R)\int_a^b f(t)\ \mathrm{d}t = \underline{\int_a^b} f(t)\ \mathrm{d}t = \overline{\int_a^b} f(t)\ \mathrm{d}t.$$

To distinguish with the Lebesgue integral, we are writing $(R)\int_a^b f(t)\ \mathrm{d}t$ to denote the Riemann integral of f.

Proposition 3.25.
Let $f \in \mathrm{Bor}_b([a,b],\mathbb{R})$. Then

$$\underline{\int_a^b} f(t)\ \mathrm{d}t \leq \int_{[a,b]} f\,\mathrm{d}\lambda_{[a,b]} \leq \overline{\int_a^b} f(t)\ \mathrm{d}t. \qquad [3.24]$$

Proof. The verifications of the two inequalities in [3.24] are similar to each other; let us see for instance the details for the first of them.

Since $\underline{\int_a^b} f(t)\ \mathrm{d}t$ is defined as a supremum of lower sums, this inequality will follow if we can verify that

$$L(f,\Delta) \leq \int_{[a,b]} f\,\mathrm{d}\lambda_{[a,b]} \qquad [3.25]$$

for all division Δ of $[a,b]$.

So let us fix a division $\Delta = (t_0,\ldots,t_k)$ of $[a,b]$, for which we will verify that [3.25] holds. For every $i \in \{1,\ldots,k\}$, we denote

$$c_i = \inf_{t_{i-1}\leq t\leq t_i} f(t).$$

This allows us to write explicitly that

$$L(f,\Delta) = \sum_{i=1}^{k} c_i\,(t_i - t_{i-1}).$$

Consider the piecewise constant function $g : [a,b] \to \mathbb{R}$ which is defined as follows: for all $t \in [a,b]$,

$$g(t) = \begin{cases} c_1 & \text{if } t_0 \leq t < t_1 \\ \vdots & \\ c_{k-1} & \text{if } t_{k-2} \leq t < t_{k-1} \\ c_k & \text{if } t_{k-1} \leq t \leq t_k \end{cases}.$$

It is immediate that $g \in \mathrm{Bor}([a,b],\mathbb{R})$, and upon comparing the definition of g with the definition of $c_1,\ldots,c_k$, we see that $g \leq f$. Moreover, g is a linear combination of indicator functions of intervals:

$$g = c_1\chi_{[t_0,t_1)} + \cdots + c_{k-1}\chi_{[t_{k-2},t_{k-1})} + c_k\chi_{[t_{k-1},t_k]}.$$

From the Lebesgue integral of g we thus have

$$\begin{aligned}\int_{[a,b]} g \, \mathrm{d}\lambda_{[a,b]} &= c_1\lambda_{[a,b]}\left([t_0, t_1)\right) + \cdots + c_{k-1}\lambda_{[a,b]}\lambda_{[a,b]}\left([t_{k-2}, t_{k-1})\right) + c_k\lambda_{[a,b]}\left([t_{k-1}, t_k]\right) \\ &= \sum_{i=1}^{k} c_i \left(t_i - t_{i-1}\right) = L\left(f, \Delta\right).\end{aligned}$$

But since Lebesgue integral is increasing, we thus have that

$$\int_{[a,b]} f \, \mathrm{d}\lambda_{[a,b]} \geq \int_{[a,b]} g \, \mathrm{d}\lambda_{[a,b]} = L\left(f, \Delta\right),$$

which is exactly the inequality that we desired to prove.

QED

Based on Proposition 3.25 and by taking into account how the Riemann integral is defined, it becomes clear that we have the following corollary.

Corollary 3.25.1.

Let $f \in \mathrm{Bor}_b\left([a, b], \mathbb{R}\right)$ and suppose that f is Riemann integrable. Then we have

$$(R) \int_a^b f(t) \, \mathrm{d}t = \int_{[a,b]} f \, \mathrm{d}\lambda_{[a,b]}.$$

Follows Immediately from Propositoin 3.25

The upshot of Corollary 3.25.1 is that, based on it, we can compute Lebesgue integrals by using the various integration techniques that we have from calculus, in reference to the Riemann integral.

IV. Hilbert Spaces and Fourier Analysis

1. $\mathcal{L}^p(\mu)$ Spaces and the Minkowski Inequality

Notation 4.1. $\mathcal{L}^p(\mu)$
Given $p \in (0,\infty)$, we denote

$$\mathcal{L}^p(\mu) = \left\{ f \in \mathrm{Bor}(X,\mathbb{R}) : \int_X |f|^p < \infty \right\}.$$

Proposition 4.1.
Let $p \in (0,\infty)$. $\mathcal{L}^p(\mu)$ is a linear subspace of $\mathrm{Bor}(X,\mathbb{R})$.

Proof. Claerly $\underline{0} \in \mathcal{L}^p(\mu)$.
Let $f, g \in \mathcal{L}^p(\mu), \alpha \in \mathbb{R}$. Then

$$\mathrm{L}^+\left(|\alpha f|^p\right) = \mathrm{L}^+\left(|\alpha|^p |f|^p\right) = |\alpha|^p \, \mathrm{L}^+\left(|f|^p\right) \infty.$$

Moreover,

$$\begin{aligned} \mathrm{L}^+\left(|f+g|^p\right) &= \int_X |f(x)+g(x)|^p \, \mathrm{d}\mu = \int_X (2\max\{|f(x)|, |g(x)|\})^p \, \mathrm{d}\mu \\ &= \int_X \max\left\{2^p |f(x)|^p, 2^p |g(x)|^p\right\} \mathrm{d}\mu \le 2^p \int_X |f(x)|^p + |g(x)|^p \, \mathrm{d}\mu = 2^p \left(\mathrm{L}^+\left(|f|^p\right) + \mathrm{L}^+\left(|g|^p\right)\right) < \infty. \end{aligned}$$

QED

As a special case, consider where $p = 1$. Then note that $\mathcal{L}^1(\mu)$ in Notation 4.1 coincides with our previous definition of $\mathcal{L}^1(\mu)$. For any $f \in \mathcal{L}^1(\mu)$, recall that we defined $\|f\|_1$ by

$$\|f\|_1 = \int_X |f| \, \mathrm{d}\mu \in [0,\infty).$$

This gave us a map $\|\cdot\|_1 : \mathcal{L}^1(\mu) \to [0,\infty)$, which is a *seminorm*. That is,

$$\begin{aligned} &\forall f, g \in \mathcal{L}^1(\mu)\left[\|f+g\|_1 \le \|f\|_1 + \|g\|_1\right], && \textit{triangle ineqautliy} \\ &\forall f \in \mathcal{L}^1(\mu), c \in \mathbb{R}\left[\|cf\|_1 = |c| \, \|f\|_1\right], && \textit{absolute homogeneity} \\ &\forall f \in \mathcal{L}^1(\mu)\left[\|f\|_1 \ge 0\right]. && \textit{nonnegativity} \end{aligned}$$

If we have in addition that

$$\forall f \in \mathcal{L}^1(\mu)\left[f \ne \underline{0} \implies \|f\|_1 > 0\right], \qquad \textit{positive definiteness}$$

then $\|\cdot\|_1$ would be a *norm*. Unfortunately, this is not the case in general.

Notation 4.2. $\|\cdot\|_p$
Let $p \in (0,\infty)$. For every $f \in \mathcal{L}^p(\mu)$, we define

$$\|f\|_p = \left(\int_X |f|^p \, \mathrm{d}\mu\right)^{\frac{1}{p}} \in [0,\infty).$$ [1]

[1] We are taking the power $\frac{1}{p}$ of the integral so that we can ensure the absolute homogeneity of $\|\cdot\|_p$.

For what $p \in (0,\infty)$ is $\|\cdot\|_p$ a seminorm? We know that for $p = 1$, $\|\cdot\|_1$ is a seminorm.

Now suppose $p \in (0, \infty) \setminus \{1\}$. It is immediate from the definition that $\|\cdot\|_p$ is absolutely homogeneous. This means it suffices to check the triangle inequality for $\|\cdot\|_p$. It turns out that the answer is positive if and only if $p > 1$. In case $p > 1$, we call the special case of the triangle inequality *Minkowski's inequality*.

Proposition 4.2. Minkowski's Inequality

Let $p > 1$. Then

$$\|f + g\|_p \le \|f\|_p + \|g\|_p$$

for all $f, g \in \mathcal{L}^p(\mu)$.

Postponed

For $p > 1$, we define a *conjugate exponent* $q \in (1, \infty)$, via requirement that $\frac{1}{p} + \frac{1}{q} = 1$. That is,

$$q = \frac{p}{p-1}.$$

There are few other equivalent definitions of q:

$$\begin{aligned} p + q &= pq, \\ q(p-1) &= p, \\ p - \frac{p}{q} &= 1. \end{aligned}$$

The idea for proving Minkowski's inequality is by playing things about $\mathcal{L}^p(\mu)$ against $\mathcal{L}^q(\mu)$ (i.e. duality). Concretely, we have the following proposition.

Proposition 4.3. Holder's Inequality

Let $p, q \in (1, \infty)$ be such that $\frac{1}{p} + \frac{1}{q} = 1$.

(a) If $f \in \mathcal{L}^p(\mu), g \in \mathcal{L}^q(\mu)$, then $fg \in \mathcal{L}^1(\mu)$ with

$$\|fg\|_1 \le \|f\|_p \|g\|_q. \qquad \textit{Holder's inequality}$$

(b) For any $f \in \mathcal{L}^p(\mu)$, we have

$$\|f\|_p = \sup\left\{\|fh\|_1 : h \in \mathcal{L}^q(\mu), \|h\|_q \le 1\right\}. \tag{4.1}$$

Postponed

It is immediate from Holder's inequality that, if $h \in \mathcal{L}^q(\mu)$ with $\|h\|_q \le 1$, then

$$\|fh\|_1 \le \|f\|_p \|h\|_q \le \|f\|_p.$$

By remembering the definition of supremum, we obtain

$$\|f\|_p \ge \sup\left\{\|fh\|_1 : h \in \mathcal{L}^q(\mu), \|h\|_q \le 1\right\}.$$

Therefore, the point of [4.1] is that we can find $h \in \mathcal{L}^q(\mu)$ to also obtain

$$\|f\|_p \le \|fh\|_1.$$

We show that Minkowski's inequality follows easily from Proposition 4.3.

Proof of Minkowski's Inequality Assuming Proposition 4.3

Let q be the conjugate exponent of p,

$$q = \frac{p}{p-1}.$$

Use [4.1] for $f+g$ to obtain

$$\|f+g\|_p = \sup\left\{\|(f+g)\,h\|_1 : h \in \mathcal{L}^q(\mu), \|h\|_q \le 1\right\}.$$

By the definition of supremum, in order to get Minkowski's inequality, it suffices to check that

$$\|(f+g)\,h\|_1 \le \|f\|_p + \|g\|_p \tag{4.2}$$

for all $h \in \mathcal{L}^q(\mu)$ with $\|h\|_q \le 1$. Hence fix $h \in \mathcal{L}^q(\mu)$ with $\|h\|_q \le 1$, for which we verify [4.2]. Indeed,

$$\begin{aligned}
\|(f+g)\,h\|_1 &= \|fh+gh\|_1 \\
&\le \|fh\|_1 + \|gh\|_1 && \text{triangle inequality for } \|\cdot\|_1 \\
&\le \|f\|_p\,\|h\|_q + \|g\|_p\,\|h\|_q && \text{Holder's inequality} \\
&\le \|f\|_p + \|g\|_q .
\end{aligned}$$

QED

We now turn to the proof of Holder's inequality.

Lemma 4.4.
Given $p, q \in (1,\infty)$ such that $\frac{1}{p}+\frac{1}{q}=1$, we have

$$\forall a, b \in [0,\infty)\left[ab \le \frac{1}{p}a^p + \frac{1}{q}b^q\right]. \tag{4.3}$$

Proof. When $a=0$ or $b=0$, [4.3] is clear. Assume $a,b>0$ and let $\alpha=\log(a), \beta=\log(b)$. Then the verification [4.3] amounts to

$$e^{\alpha}e^{\beta} \le \frac{1}{p}e^{\alpha p} + \frac{1}{q}e^{\beta q}. \tag{4.4}$$

But recall $\frac{1}{p}+\frac{1}{q}=1$, so the right-hand side of [4.4] is taking a convex combination of $e^{\alpha p}, e^{\alpha q}$. We also know that exponential functions are concave up. That is, given any $s,t\in\mathbb{R}$, $\lambda\in[0,1]$, we have

$$e^{\lambda s+(1-\lambda t)t} \le \lambda e^s + (1-\lambda)\,e^t. \tag{4.5}$$

Now by letting $\lambda = \frac{1}{p}$, we have $1-\lambda=\frac{1}{q}$ freely by definition of p,q. Then by writing $s=\alpha p, t=\beta q$, we see that [4.4] is precisely of the form [4.5], as needed.

QED

Lemma 4.5.
Let $p,q\in(1,\infty)$ with $\frac{1}{p}+\frac{1}{q}=1$ and let $f\in\mathcal{L}^p(\mu), g\in\mathcal{L}^q(\mu)$. Then $fg\in\mathcal{L}^1(\mu)$, with

$$\|fg\|_1 \le \frac{1}{p}\|f\|_p^p + \frac{1}{q}\|g\|_q^q. \tag{4.6}$$

Proof. To get [4.6] from [4.3], we take $a=|f(x)|, b=|g(x)|$ so that

$$|f(x)\,g(x)| \le \frac{1}{p}|f(x)|^p + \frac{1}{q}|g(x)|^q \tag{4.7}$$

for all $x\in X$. By integrating both sides of [4.7], we obtain [4.6].

QED

Lemma 4.6.
Let $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$ and let $f \in \mathcal{L}^p(\mu), g \in \mathcal{L}^q(\mu)$ with $\|f\|_p = \|g\|_q = 1$. Then $fg \in \mathcal{L}^1(\mu)$ with

$$\|fg\|_1 \leq 1. \tag{4.8}$$

Proof. This is a special case of Lemma 4.5, as we have

$$\|fg\|_1 \leq \frac{1}{p}\|f\|_p^p + \frac{1}{q}\|g\|_q^q = \frac{1}{p} + \frac{1}{q} = 1.$$

QED

Note that Lemma 4.6 is a special case of Holder's inequality: it says

$$\|fg\|_1 \leq 1 = \|f\|_p \|g\|_q. \tag{4.9}$$

The general statement of Holder's inequality will follow from [4.9]. The intuition is to *rescale* f, g to have norm 1, apply the special case [4.9], and then get back Holder's inequality.

The following is the conclusion of this subsection.

Proposition 4.7.
For every $p \in [1, \infty)$, $\|\cdot\|_p$ is a seminorm on $\mathcal{L}^p(\mu)$.

See Minkowski's Inequality

2. Null-space of a Seminorm

Def'n 4.3. **Null-space** of a Seminorm
Let V be a vector space over $\mathbb{R}$ and let $\|\cdot\| : V \to [0, \infty)$ be a seminorm on V. We call

$$N = \{v \in V : \|v\| = 0\}$$

the ***null-space*** of $\|\cdot\|$.

Consider the above setting. Then N is a linear subspace of V. That is,

$$\begin{aligned} &\forall v_1, v_2 \in N [v_1 + v_2 \in N], && \textit{closure under addition} \\ &\forall v \in N \forall \alpha \in \mathbb{R} [\alpha v \in N], && \textit{closure under scalar multiplication} \\ &0_V \in N. \end{aligned}$$

This means

$$\|\cdot\| \text{ is a norm on } V \iff N = \{0_V\}. \tag{4.10}$$

We now consider the space $Q = V/N$, the quotient of V by N.

Lemma 4.8.
$\|\cdot\|_Q : Q \to [0, \infty)$ defined by

$$\|\xi\|_Q = \|v\|$$

for all $\xi \in Q$, where $v \in V$ is such that $v + N = \xi$, is a norm.

Proof. To show that $\|\cdot\|_Q$ is well-defined, let $v, u \in V$ be such that $v + N = u + N$. This means $v = u + n$ for some $n \in N$, so that

$$\|v\| = \|u + n\| \leq \|u\| + \|n\| = \|u\|$$

and that

$$\|u\| = \|v - n\| \leq \|v\| + \|-n\| = \|v\|.$$

This means $\|v\| = \|u\|$, so $\|\cdot\|_Q$ is well-defined.

Given any $\xi = v + N, \eta = u + N$, we have

$$\|\xi + \eta\|_Q = \|v + u\| \leq \|v\| + \|u\| = \|\xi\|_Q + \|\eta\|_Q.$$

Moreover, given any $\alpha \in \mathbb{R}$, $\alpha\xi = \alpha(v + N) = \alpha v + N$, so that

$$\|\alpha\xi\|_Q = \|\alpha v\| = |\alpha| \|v\| = |\alpha| \|\xi\|_Q.$$

Finally, suppose $\|\xi\|_Q = 0$. This means

$$\|v\| = \|\xi\|_Q = 0,$$

so that $v \in N$. But then $v + N = 0 + N = 0_Q$, as needed.

Thus $\|\cdot\|_Q$ is a norm on Q, as required.

QED

3. The Space $\mathrm{L}^p(\mu)$

We fix a measure space $(X, \mathcal{A}, \mu)$. Recall that for every $p \in [1, \infty)$ we have a vector space $\mathcal{L}^p(\mu)$ endowed with seminorm $\|\cdot\|_p$.

Notation 4.4. $\mathcal{N}$

We are going to denote

$$\mathcal{N} = \{f \in \mathrm{Bor}(X, \mathbb{R}) : f = 0 \text{ a.e.-}\mu\}.$$

Proposition 4.9.

For every $p \in [1, \infty)$, $\mathcal{N}$ is the null-space of $\|\cdot\|_p$.

Notation 4.5. $\mathrm{L}^p(\mu)$

We write $\mathrm{L}^p(\mu)$ to denote $\mathcal{L}^p(\mu) / \mathcal{N}$.

Proposition 4.10.

Let $p \in [1, \infty)$. Consider defining $\|\cdot\|_p : \mathrm{L}^p(\mu) \to [0, \infty)$ by

$$\|f + \mathcal{N}\|_p = \|f\|_p,$$

where $\|\cdot\|_p$ on the right-hand side is the seminorm on $\mathcal{L}^p(\mu)$.[1] Then $\|\cdot\|_p$ is a norm on $\mathrm{L}^p(\mu)$.

[1] This is an unfortunate notation, of course.

Exercise 4.1.

Consider the measure space $\left([0,1], \mathcal{B}_{[0,1]}, \lambda_{[0,1]}\right)$ and $\mathcal{L}^1(\mu)$. Find $f : \mathcal{L}^1(\mu)$ with $\|f\|_1 = 0$ but $f \neq \underline{0}$.

Answer. Consider the function

$$f = \chi_{\mathbb{Q} \cap [0,1]}.$$

Then observe that $|f| = f$, where

$$\int_{[0,1]} f \,\mathrm{d}\lambda_{[0,1]} = \int_{[0,1]} \chi_{\mathbb{Q} \cap [0,1]} \,\mathrm{d}\lambda_{[0,1]} = \lambda_{[0,1]}(\mathbb{Q} \cap [0,1]) = 0.$$

Thus $f \in \mathcal{L}^1(\mu)$ with $\|f\|_1 = 0$ but $f \neq \underline{0}$.

QED

We will prove that $\left(\mathrm{L}^p(\mu), \|\cdot\|_p\right)$ is complete, hence is a Banach space. Most of the arguments will be made for $\left(\mathcal{L}^p(\mu), \|\cdot\|_p\right)$, and the last step of taking quotient by $\mathcal{N}$ will take us to the desired conclusion about $\left(\mathrm{L}^p(\mu), \|\cdot\|_p\right)$.

Since we will be working with a *semi*normed vector space, we have to adjust the notion of completeness a bit: *how does completeness work in the framework of a seminormed vector space?* It turns out that the notion of completeness works fine, but we have to be careful about limits.

To discuss things in full generality, let $(V, \|\cdot\|)$ be a seminormed vector space over $\mathbb{R}$ throughout this subsection. We do *not* assume $\|\cdot\|$ to be a norm. This means

$$\mathcal{N} = \{v \in V : \|v\| \neq 0\}$$

is a nontrivial subspace of V.

Def'n 4.6. **Convergence** of a Sequence on a Seminormed Space
Let $v \in V$ and let $(v_n)_{n=1}^{\infty} \in V^{\mathbb{N}}$. We say $(v_n)_{n=1}^{\infty}$ ***converges*** to v, denoted as $v_n \to v$, to mean that

$$\lim_{n\to\infty} \|v_n - v\| = 0.$$

Observe that the limit may not be unique. That is, suppose $(v_n)_{n=1}^{\infty} \in V^{\mathbb{N}}$ converges to $v \in V$. Since $(V, \|\cdot\|)$ is a seminormed vector space, there exists $z \in \mathcal{N}$ with $z \neq 0$. Then observe that

$$\|v_n - (v+z)\| \leq \|v_n - v\| + \|z\| = \|v_n - v\|$$

and that

$$\|v_n - v\| \leq \|v_n - (v+z)\| + \|z\| = \|v_n - (v-z)\|$$

for all $n \in \mathbb{N}$. This means $\lim_{n\to\infty} \|v_n - (v+z)\| = 0$, so that $(v_n)_{n=1}^{\infty}$ converges to $v+z$ as well.

Def'n 4.7. **Cauchy** Sequence on a Seminormed Space
Let $(v_n)_{n=1}^{\infty} \in V^{\mathbb{N}}$. We say $(v_n)_{n=1}^{\infty}$ is ***Cauchy*** if

$$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N}\, \forall n, m \geq n_0 \left[\|v_n - v_m\| < \varepsilon\right].$$

Def'n 4.8. $\frac{1}{2}$**-geometric** Sequence on a Seminormed Space
Let $(v_n)_{n=1}^{\infty} \in V^{\mathbb{N}}$. We say $(v_n)_{n=1}^{\infty}$ is $\frac{1}{2}$***-geometric*** if

$$\forall n \in \mathbb{N}\left[\|v_n - v_{n-1}\| < \frac{1}{2^n}\right].$$

Given an $\frac{1}{2}$-geometric sequence $(v_n)_{n=1}^{\infty} \in V^{\mathbb{N}}$ and $n, m \in \mathbb{N}$ with $m < n$, we have

$$\|v_m - v_n\| = \|(v_m - v_{m+1}) + (v_{m+1} - v_{m+2}) + \cdots + (v_{n-1} - v_n)\| < \sum_{j=m+1}^{n} \frac{1}{2^j}.$$

Given $(v_n)_{n=1}^{\infty} \in V^{\mathbb{N}}$, we have the following implications:

$$(v_n)_{n=1}^{\infty} \text{ is convergent} \implies (v_n)_{n=1}^{\infty} \text{ is Cauchy} \tag{4.11}$$

and that

$$(v_n)_{n=1}^{\infty} \text{ is } \frac{1}{2}\text{-geometric} \implies (v_n)_{n=1}^{\infty} \text{ is Cauchy.} \tag{4.12}$$

A weak converse of [4.12] is that

$$(v_n)_{n=1}^{\infty} \text{ is Cauchy} \implies \text{there are } n_1 < n_2 < \cdots \text{ in } \mathbb{N} \text{ such that } \left(v_{n_j}\right)_{j=1}^{\infty} \text{ is } \frac{1}{2}\text{-geometric.} \tag{4.13}$$

The converse of [4.11] is the definition of completeness.

Def'n 4.9. **Complete** Seminormed Space
We say $(V, \|\cdot\|)$ is ***complete*** if for every Cauchy sequence $(v_n)_{n=1}^{\infty} \in V^{\mathbb{N}}$, there exists $v \in V$ such that $v_n \to v$.

Now we fix a measure space $(X, \mathcal{A}, \mu)$ and $p \in [1, \infty)$, and we look at the seminormed vector space $\left(\mathcal{L}^p(\mu), \|\cdot\|_p\right)$. We will prove that $\left(\mathcal{L}^p(\mu), \|\cdot\|_p\right)$ is complete. In view of what we sah in [4.13], it will be sufficient to prove that every $\frac{1}{2}$-geometric sequence with respect to $\|\cdot\|_p$ is convergent. We start towards that by recording a general *convergence mechanism* which can be used in this kind of situation.

Lemma 4.11.
Let X be a nonempty set and let $(f_n)_{n=1}^{\infty} \in \left(\mathbb{R}^X\right)^{\mathbb{N}}$. For each $n \in \mathbb{N}$, let $h_n : X \to [0, \infty)$ be defined as

$$h_n = \sum_{m=1}^{n} |f_m - f_{m-1}|$$

with the convention that $f_0 = 0$. Denote

$$Z = \left\{x \in X : \lim_{n\to\infty} h_n(x) = \infty\right\}.$$

Then for every $x \in X \setminus Z$, $(f_n(x))_{n=1}^{\infty}$ is convergent in $\mathbb{R}$.

Proof. We pick $x \in X \setminus Z$ and verify that $(f_n(x))_{n=1}^{\infty}$ is convergent. As is always the case with sequences of real numbers, it will be sufficient to verify that

$$(f_n(x))_{n=1}^{\infty} \textit{ is a Cauchy sequence.} \tag{4.14}$$

So in connection to x picked above, suppose we are also given $\varepsilon > 0$. We want to find $n_0 \in \mathbb{N}$ such that

$$\textit{for all } n > m \geq n_0\textit{, we have } |f_m(x) - f_n(x)| < \varepsilon. \tag{4.15}$$

Since x is in $X \setminus Z$, by definition of h_n's, $(h_n(x))_{n=1}^{\infty}$ is an increasing sequence that has a finite limit, say $l \in [0, \infty)$. In connection to ε that was given, we have $n_0 \in \mathbb{N}$ such that $h_{n_0}(x) > l - \varepsilon$. We show that this n_0 does the job for [4.15]:

$$\begin{aligned}
|f_m(x) - f_n(x)| &= |(f_m(x) - f_{m+1}(x)) + (f_{m+1}(x)) - f_{m+2}(x) + \cdots + (f_{n-1}(x) - f_n(x))| \\
&\leq |f_m(x) - f_{m+1}(x)| + |f_{m+1}(x) - f_{m+2}(x)| + \cdots + |f_{n-1}(x) - f_n(x)| \\
&= h_n(x) - h_M(x) \\
&< l - (l - \varepsilon) && \text{since } h_n(x) \leq l \text{ and } h_m(x) \geq h_{n_0}(x) > l - \varepsilon \\
&= \varepsilon.
\end{aligned}$$

Thus $|f_m(x) - f_n(x)| < \varepsilon$, as required.

QED

Proposition 4.12.
Let $(f_n)_{n=1}^{\infty} \in \mathcal{L}^p(\mu)^{\mathbb{N}}$ be such that $\|f_n - f_{n+1}\|_p < \frac{1}{2^n}$ for all $n \in \mathbb{N}$. Consier Z from Lemma 4.11 and define $f : X \to \mathbb{R}$ by

$$f(x) = \begin{cases} 0 & \text{if } x \in Z \\ \lim_{n\to\infty} f_n(x) & \text{if } x \in X \setminus Z \end{cases}$$

for all $x \in X$. Then $f \in \mathcal{L}^p(\mu)$ with $\lim_{n\to\infty} \|f_n - f\|_p = 0$.

tl;dr

Corollary 4.12.1.
The seminormed vector space $\left(\mathcal{L}^p(\mu), \|\cdot\|_p\right)$ is complete.

Corollary 4.12.2.
The normed vector space $\left(\mathrm{L}^p(\mu), \|\cdot\|_p\right)$ is a Banach space.

A byproduct of how we ran the proof of Proposition 4.12 is that we can connect convergence with respect to $\|\cdot\|_p$ to a notion of *almost everywhere* convergence, which is defined as follows.

Def'n 4.10. **Converges Almost Everywhere**
Let $(X, \mathcal{A}, \mu)$ be a measure space, let $f \in \operatorname{Bor}(X, \mathbb{R})$, and let $(f_n)_{n=1}^{\infty} \in \operatorname{Bor}(X, \mathbb{R})^{\mathbb{N}}$. We say $(f_n)_{n=1}^{\infty}$ ***converges*** to f ***almost everywhere*** with respect to μ (or ***a.e.***-μ for short) to mean that there exists a negligible set $Z \in \mathcal{A}$ such that

$$\lim_{n\to\infty} f_n(x) = f(x)$$

for all $x \in X \setminus Z$.

The relevance of Def'n 4.10 comes up as follows. Upon examining the proof of Proposition of 4.12, it is quite obvious that the function f considered there also has the property that $(f_n)_{n=1}^{\infty}$ converges to f in the a.e.-μ sense. Indeed, the convergence $f_n(x) \to f(x)$ may only fail for points x in the set Z, and $\mu(Z) = 0$ as in the proof. We can therefore also record the following corollary.

Corollary 4.12.3.
Let $(X, \mathcal{A}, \mu)$ be a measure space, let $p \in [1, \infty)$, and let $(f_n)_{n=1}^{\infty} \in \mathcal{L}^p(\mu)^{\mathbb{N}}$ be such that $\|f_n - f_{n+1}\|_p < \frac{1}{2^n}$. Then there exists $f \in \mathcal{L}^p(\mu)$ such that $\lim_{n\to\infty} f_n = f$ a.e.-μ.

5. Hilbert Spaces

Recall 4.11. **Inner Product Space** over $\mathbb{R}$
Let V be an $\mathbb{R}$-vector space. We say $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ is an ***inner product*** on V if

(a) for all $u, v, w \in V, a \in \mathbb{R}$, we have $\langle au + v, w\rangle = a\langle u, w\rangle + \langle v, w\rangle$; *linearity at the first argument*
(b) for all $u, v \in V$, we have $\langle u, v\rangle = \langle v, u\rangle$; and *symmetry*
(c) for all nonzero $v \in V$, we have $\langle v, v\rangle > 0$. *positive definiteness*

Whenever $\langle\cdot,\cdot\rangle$ is an inner product on V, we say $(V, \langle\cdot,\cdot\rangle)$ is an ***inner product space***.

Observe that, for any $v \in V$,

$$\langle 0_V, v\rangle = \langle 0v, v\rangle = 0\langle v, v\rangle = 0$$

by using the linearity at the first argument.

Proposition 4.13. Cauchy-Schwarz Inequality
Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space. Then

$$\langle v, w\rangle^2 \le \langle v, v\rangle\langle w, w\rangle$$

for all $v, w \in V$.

MATH 245

Proposition 4.14.
Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space. Then $\|\cdot\| : V \to \mathbb{R}$ by

$$\|v\| = \sqrt{\langle v, v\rangle}$$

is a norm on V.

Recall 4.12. **Norm** Induced by the Inner Product
Consider the setting of Proposition 4.14. We call $\|\cdot\|$ the ***norm*** induced by $\langle\cdot,\cdot\rangle$.

Def'n 4.13. **Hilbert Space**
Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space and let $\|\cdot\|$ be the norm induced by $\langle\cdot,\cdot\rangle$. If $\|\cdot\|$ induces a complete metric on V, then we say $\langle V, \langle\cdot,\cdot\rangle\rangle$ is a ***Hilbert space***.

The space $\mathrm{L}^2(\mu)$ is special among $\mathrm{L}^p(\mu)$ spaces in the sense that the conjugate exponent q of $p = 2$ requires $\frac{1}{p} + \frac{1}{q} = 1$, which means $q = p = 2$. So then Holder's inequality says, given $f, g \in \mathcal{L}^2(\mu), fg \in \mathcal{L}^1(\mu)$ and

$$\|fg\|_1 \leq \|f\|_2 \|g\|_2 .$$

This suggests an inner product space structure on $\mathrm{L}^2(\mu)$.

Proposition 4.15.
The map $\langle\cdot,\cdot\rangle : \mathrm{L}^2(\mu) \times \mathrm{L}^2(\mu) \to \mathbb{R}$ by

$$\langle f+\mathcal{N}, g+\mathcal{N}\rangle = \int_X fg \,\mathrm{d}\mu$$

for all $f+\mathcal{N}, g+\mathcal{N} \in \mathrm{L}^2(\mu)$ is a complete inner product on $\mathrm{L}^2(\mu)$. Moreover, the norm induced by $\langle\cdot,\cdot\rangle$ is $\|\cdot\|_2$.

Proof. Given any $f+\mathcal{N}, g+\mathcal{N} \in \mathrm{L}^2(\mu), f, g \in \mathcal{L}^2(\mu)$ so that $fg \in \mathcal{L}^1(\mu)$ by Holder's inequality. This means $\int_X fg \,\mathrm{d}\mu$ is well-defined.

Moreover, if $f_1, g_1 \in \mathcal{L}^2(\mu)$ are such that $f+\mathcal{N} = f_1+\mathcal{N}$ and $g+\mathcal{N} = g_1+\mathcal{N}$, $fg = f_1g_1$ a.e.-μ, so that $\langle f+\mathcal{N}, g+\mathcal{N}\rangle = \langle f_1+\mathcal{N}, g_1+\mathcal{N}\rangle$. Hence $\langle\cdot,\cdot\rangle$ is well-defined.

Now let $f, g, h \in \mathcal{L}^2(\mu), a \in \mathbb{R}$. Then

$$\begin{aligned}
\langle af+g+\mathcal{N}, h+\mathcal{N}\rangle &= \int_X (af+g)h \,\mathrm{d}\mu = a\int_X fh \,\mathrm{d}\mu + \int_X gh \,\mathrm{d}\mu = a\langle f+\mathcal{N}, h+\mathcal{N}\rangle + \langle g+\mathcal{N}, h+\mathcal{N}\rangle, \\
\langle f+\mathcal{N}, g+\mathcal{N}\rangle &= \int_X fg \,\mathrm{d}\mu = \int_X gf \,\mathrm{d}\mu = \langle g, f\rangle, \\
\langle f+\mathcal{N}, f+\mathcal{N}\rangle &= \int_X f^2 \,\mathrm{d}\mu = \|f\|_2^2 = \|f+\mathcal{N}\|_2^2,
\end{aligned}$$

where the last line shows $\langle\cdot,\cdot\rangle$ is positive definite. It then shows that $\|\cdot\|_2$ is induced by $\langle\cdot,\cdot\rangle$. But we know that $\|\cdot\|_2$ is complete (Corollary 4.12.1). Thus $(\mathrm{L}^2(\mu), \langle\cdot,\cdot\rangle)$ is a Hilbert space.

QED

6. Geometry of Hilbert Spaces

Recall 4.14. **Convex** Subset of an $\mathbb{R}$-vector Space
Let V be an $\mathbb{R}$-vector space. Given any $u, v \in V$, we define the ***line segment*** with endpoints u, v, denoted as $\mathrm{Co}(u, v)$, by

$$\mathrm{Co}(u, v) = \{(1-t)v + tw : t \in [0,1]\}.$$

If $C \subseteq V$ is such that, for all $u, v \in C$, $\mathrm{Co}(u, v)$ is contained in C, then we say C is ***convex***.

Exercise 4.2.
Let $(V, \|\cdot\|)$ be a normed $\mathbb{R}$-vector space. Prove that, if $C \subseteq V$ is convex, then $\mathrm{cl}(C)$ is also convex.

Proposition 4.16.
Let $(V, \langle\cdot,\cdot\rangle)$ be a Hilbert space and let $C \subseteq V$ be a nonempty closed convex set. Then for all $v \in V \setminus C$, there exists unique $w_0 \in C$ such that

$$\|v - w_0\| = \inf_{w\in C} \|v - w\|.$$

Proof. Suppose $v \notin C$ (otherwise, we may set $w_0 = v$). Denote

$$\alpha = \inf_{w \in C} \|v - w\| .$$

We claim that

$$\alpha > 0. \tag{4.16}$$

For contradiction, suppose that $\alpha = 0$. Then for every $n \in \mathbb{N}$, there is $w_n \in C$ such that $\|v - w_n\| < \frac{1}{n}$. This means

$$\lim_{n\to\infty} \|v - w_n\| = \lim_{n\to\infty} \frac{1}{n} = 0.$$

But $\|\cdot\|$ is a norm on V, so it follows that $v = \lim_{n\to\infty} w_n \in \operatorname{cl}(C) = C$. This contradicts the assumption that $v \notin C$. Hence [4.16] is verified.

We now have to show that

$$\textit{there exists unique } w_0 \in C \textit{ such that } \|v - w_0\| = \alpha. \tag{4.17}$$

From the definition of α as an infimum, we see that, for all $n \in \mathbb{N}$ there exists $w_n \in \varphi$ such that

$$\alpha \le \|v - w_n\| < \alpha + \frac{1}{n}.^{1} \tag{4.18}$$

We claim that

$$\forall m, n \in \mathbb{N}\left[\|w_m - w_n\|^2 \le 2\left(\frac{2\alpha}{m} + \frac{2\alpha}{n} + \frac{1}{m^2} + \frac{1}{n^2}\right)\right]. \tag{4.19}$$

To prove [4.19], we will use the *parallelogram law*:

$$\forall x, y \in V\left[\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)\right]. \tag{4.20}$$

Observe that, for all $m, n \in \mathbb{N}$,

$$\begin{aligned}
\|w_m - w_n\|^2 &= \|(w_m - v) - (w_n - v)\|^2 \\
&= 2\left(\|w_m - v\|^2 + \|w_n - v\|^2\right) - \|(w_m - v) + (w_n - v)\|^2 && \text{by [4.20]} \\
&\le 2\left(\left(\alpha + \frac{1}{m}\right)^2 + \left(\alpha + \frac{1}{n}\right)^2\right) - \left\|-2\left(v - \frac{w_m + w_n}{2}\right)\right\|^2 && \text{by [4.18]} \\
&\le 2\left(\left(\alpha + \frac{1}{m}\right)^2 + \left(\alpha + \frac{1}{n}\right)^2\right) - 4\alpha^2 \\
&= 2\left(\frac{2\alpha}{m} + \frac{1}{m^2} + \frac{2\alpha}{n} + \frac{1}{n^2}\right),
\end{aligned}$$

where we have

$$\left\|-2\left(v - \frac{w_m + w_n}{2}\right)\right\| = 2\left\|v - \frac{w_m + w_n}{2}\right\| \ge 2\alpha$$

by observing that $\frac{w_m + w_n}{2} \in C$ by the convexity of C.

As a corollary to [4.19], we have

$$(w_n)_{n=1}^{\infty} \textit{ is Cauchy.} \tag{4.21}$$

Given any $\varepsilon > 0$, choose $n_0 \in \mathbb{N}$ such that

$$2\left(\frac{4\alpha}{n_0} + \frac{2}{n_0^2}\right) < \varepsilon^2.$$

Then for all $m, n \ge n_0$, we have

$$\begin{aligned}
\|w_n - w_m\| &\le \sqrt{2\left(\frac{2\alpha}{m} + \frac{2\alpha}{n} + \frac{1}{m^2} + \frac{1}{n^2}\right)} && \text{[4.19]} \\
&\le \sqrt{2\left(\frac{4\alpha}{n_0} + \frac{2}{n_0^2}\right)} < \sqrt{\varepsilon^2} = \varepsilon.
\end{aligned}$$

Hence [4.21] is verified.

Since $(V, \langle\cdot,\cdot\rangle)$ is complete, it follows that $(w_n)_{n=1}^{\infty}$ is convergent. This means we may write

$$w_0 = \lim_{n\to\infty} w_n.$$

Note that $w_0 \in C$, since every $w_n \in C$ and C is closed. Consequently,

$$\lim_{n\to\infty} \|w_n - v\| = \|w_0 - v\| = \alpha,$$

where the last equality holds by the squeeze theorem on [4.18].

The uniqueness part of [4.17] will be on an assignment.

[1]From [4.18], it might be tempting to extract a convergent subsequence of $(w_n)_{n=1}^{\infty}$; this only works in *finite* dimensional cases.

QED

Exercise 4.3.

Let $(V, \|\cdot\|)$ be a normed vector space.

(a) Let W be a linear subspace of V. Prove that $\mathrm{cl}(W)$ is also a linear subspace of V.

(b) Let W be a finite dimensional linear subspace of V. Prove that W is closed.

Proof of (a). Since W is a linear subspace of V, $0_V \in W \subseteq \mathrm{cl}(W)$.

Let $v, w \in \mathrm{cl}(W), a \in \mathbb{R}$. We have to show $u = av + w \in \mathrm{cl}(W)$. Since $v, w \in \mathrm{cl}(W)$, there exist sequences $(v_n)_{n=1}^{\infty}, (w_n)_{n=1}^{\infty} \in W^{\mathbb{N}}$ such that $v_n \to v, w_n \to w$. Define $(u_n)_{n=1}^{\infty}$ by

$$u_n = av_n + w_n$$

for all $n \in \mathbb{N}$. Then $(u_n)_{n=1}^{\infty} \in W^{\mathbb{N}}$, since W is closed under linear combination. Moreover, since $v_n \to v, w_n \to w$, it follows $u_n \to av + w = u$. Hence $u \in \mathrm{cl}(W)$.

Proof of (b). Let $\{w_1, \ldots, w_d\}$ be a basis for W, where $d = \dim(W)$. Let $(v_n)_{n=1}^{\infty} \in W^{\mathbb{N}}$ be convergent in V, where we have to show $(v_n)_{n=1}^{\infty}$ converges to a point in W.

For each $n \in \mathbb{N}$, write

$$v_n = \sum_{j=1}^{d} a_j^{(n)} w_j$$

for some $a_1^{(n)}, \ldots, a_d^{(n)} \in \mathbb{R}$. We claim that

$$\text{each sequence } \left(a_j^{(n)}\right)_{n=1}^{\infty} \text{ converges to a point } a_j \in \mathbb{R}. \qquad [4.22]$$

Suppose [4.22] is false, for contradiction. Then $\left(a_{j=1}^{(n)}\right)_{j=1}^{\infty}$ is not Cauchy, which implies $(v_n)_{n=1}^{\infty}$ is not Cauchy, hence not convergent.

Now observe that $v_n \to \sum_{j=1}^{d} a_j w_j$.

QED

Def'n 4.15. **Orthogonal** Subsets of an Inner Product Space

Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space. For all $v, w \in V$, we say v, w are *orthogonal* to each other, denoted as $v \perp w$, when $\langle v, w\rangle = 0$.

Given nonempty $A, B \in V$, we say A, B are *orthogonal* to each other, denoted as $A \perp B$, to mean that $v \perp w$ for all $v \in A, w \in B$.

Proposition 4.17.

Consider the setting of Proposition 4.17. If C is a closed linear subspace of V, then $v - w_0$ is orthogonal to W.

Proposition 4.17 suggests the following definition.

Def'n 4.16. **Orthogonal Projection** of a Vector onto a Closed Subspace

Let $(V, \langle \cdot, \cdot \rangle)$ be a Hilbert space and let W be a closed linear subspace of V. Let $v \in V$ and let $w_0 \in W$ be the unique vector in W, provided by Proposition 4.17 (with $C = W$), which has

$$\|v - w_0\| = \inf_{w \in W} \|v - w\| .$$

We call w_0 the ***orthogonal projection*** of v onto W.

7. Orthogonal Complements

We fix a Hilbert space $(V, \langle \cdot, \cdot \rangle)$ and put into evidence a number of useful facts related to the orthogonality relation between vectors in V.

Def'n 4.17. **Orthogonal Complement** of a Subset

Let $A \subseteq V$ be nonempty. The ***orthogonal complement*** of A, denoted as $A^\perp$, is defined as

$$A^\perp = \{z \in V : \forall v \in A \left[\langle v, z \rangle = 0\right]\} .$$

Exercise 4.4.

Prove that for every nonempty $A \subseteq V$, $A^\perp$ is a closed linear subspace of V.

Proof. Since $\langle v, 0_V \rangle = 0$ for all $v \in A$, $0_V \in A^\perp$.

Let $x, y \in A^\perp$, $a \in \mathbb{R}$. We have to show $z = ax + y \in A^\perp$. Let $v \in A$. Then observe that

$$\langle v, z \rangle = \langle v, ax + y \rangle = a \langle v, x \rangle + \langle v, y \rangle = a0 + 0 = 0.$$

Hence $z \in A^\perp$.

Let $(x_n)_{n=1}^\infty$ be a convergent sequence on $A^\perp$, where we have to show $(x_n)_{n=1}^\infty$ converges to a point in $A^\perp$. Let $x = \lim_{n\to\infty} x_n$. Then observe that, given any $v \in A$,

$$\langle v, x \rangle = \left\langle v, \lim_{n\to\infty} x_n \right\rangle = \lim_{n\to\infty} \langle v, x_n \rangle = 0$$

by using the continuity of $\langle \cdot, \cdot \rangle$.

QED

Exercise 4.5.

Let $A, B \subseteq V$ be nonempty. Prove that, if $A \subseteq B$, then $B^\perp \subseteq A^\perp$.

Proof. Let $z \in B^\perp$. Then given any $v \in A$, since $A \subseteq B$, $v \in B$, and consequently $\langle v, z \rangle = 0$. Hence $v \in A^\perp$, implying $B^\perp \subseteq A^\perp$.

QED

Def'n 4.18. **Closed Linear Span** of a Set

Let $A \subseteq V$ be nonempty. We define the ***closed linear span*** of A, denoted as cl-span (A), is defined as

$$\text{cl-span}(A) = \text{cl}(\text{span}(A)) .$$

Since the closure of a linear subspace is a closed linear subspace, it follows that cl-span (A) is a closed linear subspace of V. It is easy to see that cl-span (A) is in fact the *smallest* closed linear subspace of V that contains A.

Proposition 4.18.

Let $A \subseteq V$ be nonempty and let $W = \text{cl-span}(W)$. Then $A^\perp = W^\perp$.

Proof. Since $A \subseteq W$ by definition, we have $A^\perp \supseteq W^\perp$. Hence it suffices to show $A^\perp \subseteq W^\perp$. Let $z \in A^\perp$ and let $v \in W$. Since $v \in W = \text{cl-span}(A)$, there exists $(v_n)_{n=1}^\infty \in \text{span}(A)^{\mathbb{N}}$ that converges to v. Now, for all $n \in \mathbb{N}$, $\langle z, v_n \rangle = 0$, since v_n is a linear combination of elements of A and every element of A is orthogoanl to z. Hence it follows that

$$\langle z, v \rangle = \left\langle z, \lim_{n\to\infty} v_n \right\rangle = \lim_{n\to\infty} \langle z, v_n \rangle = 0.$$

This means $z \in W^\perp$, so that $A^\perp \subseteq W^\perp$, as needed.

QED

Exercise 4.6.

Let W be a closed linear subspace of V. Prove that $W \cap W^\perp = \{0_V\}$.

Proof. Suppose for contradiction there is $v \in W \cap W^\perp$ with $v \neq 0_V$. This means $\langle v, v \rangle = 0$, contradicting the positive definiteness of $\langle \cdot, \cdot \rangle$.

QED

Proposition 4.19.

Let W be a closed linear subspace of V and let $v \in V$. Let w_0, z_0 be the orthogonal projections of v onto W, $W^\perp$, respectively. Then

(a) $w_0 + z_0 = v$; and

(b) if there is $w \in W, z \in W^\perp$ such that $w + z = v$, then $w = w_0, z = z_0$.

Corollary 4.19.1.

Let W be a closed linear subspace of V. Then $(W^\perp)^\perp = W$.

8. Countable Orthonormal Basis

We start our discussion of Fourier coefficients and Fourier expansion for a function in $L^2(\lambda_{[-\pi,\pi]})$, where $\lambda_{[-\pi,\pi]}$ is the Lebesgue measure on the Borel σ-algebra of $[-\pi, \pi]$. The results we want to obtain are actually holding in the more general framework of a Hilbert space where we have spotted an infinite countable orthonormal basis. These results also become more transparent if we discuss them in the general Hilbert space framework; so, for the moment, the space $L^2(\lambda_{[-\pi,\pi]})$ will be disguised as a Hilbert space V. In the next lecture we will let V become the required L^2-space, and we will clarify what is the orthonormal basis that we want to work with.

Def'n 4.19. **Orthonormal Basis** of a Hilbert Space

Let $(V, \langle \cdot, \cdot \rangle)$ be a Hilbert space and let $\{\xi_i\}_{i=0}^\infty \subseteq V$. If

(a) $\langle \xi_i, \xi_j \rangle = \delta_{i,j}$ (where $\delta_{i,j}$ is the Kronecker delta) for all $i, j \in \mathbb{N} \cup \{0\}$; and *orthonormality*

(b) $\text{cl-span}(\{\xi_i\}_{i=0}^\infty) = V$; *density of span*

then we say $\{\xi_i\}_{i=0}^\infty$ is an ***orthonormal basis*** for V.

Let $(V, \langle \cdot, \cdot \rangle)$ be a Hilbert space and let $\mathcal{X} = \{\xi_i\}_{i=0}^\infty$ be an orthonormal basis for V. We easily see that vectors in $\mathcal{X}$ are distinct, for if $\xi_i = \xi_j$ for some $i \neq j$ in $\mathbb{N} \cup \{0\}$, then we have

$$0 = \delta_{i,j} = \langle \xi_i, \xi_j \rangle = \langle \xi_i, \xi_i \rangle = \delta_{i,i} = 1,$$

which is a contradiction.

There are several ways to think of the condition $\text{cl-span}(\mathcal{X}) = V$.

(a) One is to simply recall that $\text{cl-span}(\mathcal{X}) = \text{cl}(\text{span}(\mathcal{X}))$. Hence if we denote

$$S = \text{span}(\mathcal{X}) = \left\{ v \in V : \exists n \in \mathbb{N} \cup \{0\}\ \exists \alpha_0, \ldots, \alpha_n \in \mathbb{R} \left[v = \sum_{i=0}^n \alpha_i \xi_i \right] \right\}, \tag{4.23}$$

we have that $\text{cl}(S) = V$. In other words, ths span S of $\mathcal{X}$ is dense in V.

(b) A second way goes by using the interpretation of $\text{cl-span}(\mathcal{X})$ as the smallest possible closed linear subspace containing $\mathcal{X}$. The condition says that this subspace has to be, in fact, the full space V. So the condition can be rephrased as

$$\textit{if } U \textit{ is a closed linear subspace of } V \textit{ containing } \mathcal{X}, \textit{ then } U = V. \tag{4.24}$$

(c) Finally, we take advantage of the operation $\perp$. If the condition holds, then we obtain that

$$\mathcal{X}^\perp = (\text{cl-span}(\mathcal{X}))^\perp = V^\perp = \{0_V\}.$$

Conversely, if $\mathcal{X}^\perp = \{0_V\}$, then we have

$$\text{cl-span}(\mathcal{X}) = \left(\mathcal{X}^\perp\right)^\perp = \{0_V\}^\perp = V.$$

Hence the condition $\text{cl-span}(\mathcal{X}) = V$ is equivalent to $\mathcal{X}^\perp = \{0_V\}$. In words:

$$\textit{if } v \in V \textit{ is such that } \langle v, \xi_i \rangle = 0 \textit{ for all } i \in \mathbb{N} \cup \{0\}, \textit{ then } v = 0_V. \tag{4.25}$$

These observations can be strengthened to a statement about linear independence.

Proposition 4.20.

Let $(V, \langle\cdot,\cdot\rangle)$ be a Hilbert space and let $\mathcal{X} = \{\xi_i\}_{i=0}^\infty \subseteq V$ be an orthonormal basis for V.

(a) For every $n \in \mathbb{N} \cup \{0\}$, $\xi_0, \ldots, \xi_n$ are linearly independent.

(b) For every $n \in \mathbb{N} \cup \{0\}$, let us write

$$V_n = \text{span}(\xi_0, \ldots, \xi_n).$$

Then V_n is an $(n+1)$-dimensional linear subspace of V, where $\xi_0, \ldots, \xi_n$ form a basis. The norms of the vectors from V_n satisfy

$$\left\| \sum_{i=0}^n \alpha_i \xi_i \right\| = \sqrt{\sum_{i=0}^n \alpha_i^2}$$

for all $\alpha_0, \ldots, \alpha_n \in \mathbb{R}$.

(c) The subspaces $V_0, V_1, \ldots$ satisfy

$$V_0 \subseteq V_1 \subseteq \cdots$$

with

$$\text{span}(\mathcal{X}) = \bigcup_{n=0}^\infty V_n.$$

As a result, $\bigcup_{n=0}^\infty V_n$ is a linear subspace of V.

Proof of (a). Suppose $\xi_0, \ldots, \xi_n$ are not linearly independent, for contradiction. This means

$$\xi_0 = \sum_{i=1}^n \alpha_i \xi_i$$

for some nonzero $(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$. Then we have $i \in \{1, \ldots, n\}$ such that $\alpha_i \neq 0$, so that

$$0 = \langle \xi_0, \xi_i \rangle = \sum_{j=1}^n \alpha_j \left\langle \xi_j, \xi_i \right\rangle = \alpha_i \neq 0,$$

which is a contradiction.

Proof of (b). Since $\xi_0, \ldots, \xi_n$ are linearly independent, it follows $V_n = \operatorname{span}\{\xi_i\}_{i=0}^n$ is an $(n+1)$-dimensional subspace of V, with a basis $\{\xi_i\}_{i=0}^n$. Now, given any $\alpha_0, \ldots, \alpha_n \in \mathbb{R}$, we have

$$\left\|\sum_{i=0}^n \alpha_i\xi_i\right\| = \sqrt{\left\langle \sum_{i=0}^n \alpha_i\xi_i, \sum_{j=0}^n \alpha_j\xi_j\right\rangle} = \sqrt{\sum_{i,j=0}^n \alpha_i\alpha_j\left\langle \xi_i, \xi_j\right\rangle} = \sqrt{\sum_{i=0}^n \alpha_i^2}.$$

Proof of (c). Since $\{\xi_i\}_{i=0}^n \subseteq \{\xi_i\}_{i=0}^m$ for all $n \leq m$ in $\mathbb{N}\cup\{0\}$, we have $V_0 \subseteq V_1 \subseteq \cdots$. Since $\{\xi_i\}_{i=0}^n \subseteq \mathcal{X}$ for all $n \in \mathbb{N}\cup\{0\}$, $V_n = \operatorname{span}\{\xi_i\}_{i=0}^n \subseteq \operatorname{span}(\mathcal{X})$ for all $n \in \cup\mathbb{N}\{0\}$. It follows $\bigcup_{n=0}^\infty V_n \subseteq \operatorname{span}(\mathcal{X})$.

Conversely, given any $v \in \operatorname{span}(\mathcal{X})$, we have $n \in \mathbb{N}\cup\{0\}$ and $\alpha_0, \ldots, \alpha_n \in \mathbb{R}$ such that $v = \sum_{i=0}^n \alpha_i\xi_i$, so that $v \in V_n \subseteq \bigcup_{m=0}^\infty V_m$. Thus $\bigcup_{n=0}^\infty V_n = \mathcal{X}$.

QED

Let $(V, \langle\cdot,\cdot\rangle)$ be a Hilbert space and let $\mathcal{X} = \{\xi_i\}_{i=0}^\infty \subseteq V$ be an orthonormal basis for V. The existence of the orthonormal basis $\mathcal{X}$ has the consequence that the Hilbert space we are dealing with a *separable, infinite dimensional* Hilbert space. That is, the vector space V is infinite dimensional andit is separable (it admits a countable dense subset) with respect to the norm $\|\cdot\|$ associated to the inner product.

The fact that V is infinite dimensional comes as an obvious consequence of Proposition 4.20, which shows that V has finite dimensional linear subspaces of arbitrary large dimension.

In order to check that V is separable in the metric given by its norm, we make an adjustment in the nature of the coefficients α_i's used in [4.23]. That is, we put

$$S_0 = \left\{v \in V : \exists n \in \mathbb{N}\cup\{0\}\ \exists \alpha_0, \ldots, \alpha_n \in \mathbb{Q}\left[v = \sum_{i=0}^n \alpha_i\xi_i\right]\right\}.$$

It is quite straightforward to show that S_0 is a countable set. Let us also argue that S_0 is dense in V.

Lemma 4.21.
Let $v \in V$ and $\varepsilon > 0$ be given. There exists $w \in S_0$ such that $\|v - w\| < \varepsilon$.

Proof. Let $S = \operatorname{span}(\mathcal{X})$. We noticed that $\operatorname{cl}(S) = V$. In particular, we have that $v \in \operatorname{cl}(S)$, and we can therefore find $v' \in S$ such that $\|v - v'\| < \frac{\varepsilon}{2}$. Now v' can be written in the form

$$v' = \sum_{i=0}^n \alpha_i\xi_i$$

for some $n \in \mathbb{N}\cup\{0\}$ and $\alpha_0, \ldots, \alpha_n \in \mathbb{R}$. Let $r_0, \ldots, r_n \in \mathbb{Q}$ be such that $|r_i - \alpha_i| < \frac{\varepsilon}{\sqrt{4(n+1)}}$ for every $i \in \{0, \ldots, n\}$ and conside the vector

$$w = \sum_{i=0}^n r_i\xi_i \in S_0.$$

Then we have

$$\left\|v' - w\right\|^2 = \left\|\sum_{i=0}^n (\alpha_i - r_i)\,\xi_i\right\|^2 = \sum_{i=0}^n (\alpha_i - r_i)^2 < \sum_{i=0}^n \left(\frac{\varepsilon}{\sqrt{4\,(n+1)}}\right)^2 = \frac{\varepsilon^2}{4}$$

so that $\|v' - w\| < \frac{\varepsilon}{2}$. Thus

$$\|w - v\| \leq \left\|w - v'\right\| + \left\|v' - v\right\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

as required.

QED

It is not hard to verify that Lemma 4.21 has a converse, stated as follows.

Proposition 4.22.
Let $(V, \langle\cdot,\cdot\rangle)$ be a Hilbert space. If $(V, \langle\cdot,\cdot\rangle)$ is separable and infinite dimensional, then V admits an orthonormal basis.

Assignment!

It is nice to know that an orthonormal basis can always be found if we know that the Hilbert space is separable and infinite dimensional. But we will not need to rely on this fact in our incoming discussion of $L^2(\lambda_{[-\pi,\pi]})$. Indeed, in the case of $L^2(\lambda_{[-\pi,\pi]})$, we will have available an explicit construction of orthonormal basis that we want to use, so we will not need to invoke the general orthonormal basis existence result.

For the upcoming discussion, fix a Hilbert space $(V, \langle\cdot,\cdot\rangle)$ and an orthonormal basis $\mathcal{X} = \{\xi_i\}_{i=1}^{\infty} \subseteq V$ for V.

Def'n 4.20. ***i*-th Coefficient** of a Vector

Let $v \in V$. For every $i \in \mathbb{N} \cup \{0\}$, we call

$$c_i = \langle v, \xi_i \rangle$$

the ***i*-th coefficient** of v with respect to $\mathcal{X}$. In short, we will refer to $(c_i)_{i=0}^{\infty}$ the (sequence of) $\mathcal{X}$-***coefficients*** of v.

Let us make the observation that the sequence of $\mathcal{X}$-coefficients of a vector $v \in V$ will uniquely determine what v is.

Proposition 4.23.

Let $v, v' \in V$ and let $(c_i)_{i=0}^{\infty}, (c_i')_{i=0}^{\infty}$ be the sequences of $\mathcal{X}$-coefficients of v, v', respectively. If $c_i = c_i'$ for all $i \in \mathbb{N} \cup \{0\}$, then $v = v'$.

Proof. Observe that, for all $i \in \mathbb{N} \cup \{0\}$, we have

$$\left\langle v - v', \xi_i \right\rangle = \langle v, \xi_i \rangle - \left\langle v', \xi_i \right\rangle = c_i - c_i = 0.$$

It follows from [4.25] that $v - v' = 0_V$, so that $v = v'$.

QED

We next look at how a vector $v \in V$ is approximated with linear combinations created by using $\mathcal{X}$-coefficients. To this end we use the finite dimensional spaces $V_n = \operatorname{span}\{\xi_i\}_{i=0}^{n}$.

Lemma 4.24. Bessel's Inequality

Let $v \in V$ and let $(c_i)_{i=1}^{\infty}$ be the sequence of $\mathcal{X}$-coefficients of v. Then

$$\sum_{i=0}^{n} c_i^2 \le \|v\|^2$$

for all $n \in \mathbb{N} \cup \{0\}$.

Proof. Let $v_n = \sum_{i=0}^{n} c_i \xi_i$. Then, for every $i \le N$,

$$\langle v - v_n, \xi_i \rangle = \langle v, \xi_i \rangle - \langle v_n, \xi_i \rangle = c_i - c_i = 0. \qquad [4.26]$$

Since v_n is a linear combination of $\xi_0, \ldots, \xi_n$, it follows

$$\langle v_n, v - v_n \rangle = 0.$$

Thus by the Pythagorean equality,

$$\sum_{i=0}^{n} c_i^2 = \|v_n\|^2 \le \|v_n\|^2 + \|v - v_n\|^2 = \|v\|^2 .$$

QED

Proposition 4.25.

Let $v \in V$ and let $(c_i)_{i=0}^{\infty}$ be the $\mathcal{X}$-coefficients of v. For all $n \in \mathbb{N} \cup \{0\}$, define

$$v_n = \sum_{i=0}^{n} c_i \xi_i .$$

(a) For all $n \in \mathbb{N} \cup \{0\}$, v_n is the orthogonal projection of v onto $V_n = \operatorname{span}\{\xi_i\}_{i=0}^n$. That is,

$$\|v - v_n\| = \inf\left\{\left\|v - \sum_{i=0}^{n} \alpha_i \xi_i\right\| : \alpha_0, \ldots, \alpha_n \in \mathbb{R}\right\}.$$

(b) We have

$$\lim_{n\to\infty} \|v_n - v\| = 0.$$

Proof of (a). By [4.26], we have that $v - v_n \in V_n^{\perp}$. This means $v_n \in V_n, v - v_n \in V_n^{\perp}$ are such that $v_n + (v - v_n) = v$. This means v_n is the orthogonal projection of v onto V_n.

Proof of (b). By Bessel's equality and the monotone convergence theorem (for sequences), the series $\sum_{n=0}^{\infty} c_i^2$ is convergent. This means, for all $\varepsilon > 0$, there is $N \in \mathbb{N}$ such that

$$\sum_{i=N}^{\infty} c_i^2 < \varepsilon.$$

Then given any $m, n > N$ (say $n < m$ for convenience), we have

$$\|v_m - v_n\|^2 = \sum_{i=n}^{m} c_i^2 \le \sum_{i=N}^{m} c_i^2 \le \sum_{i=N}^{\infty} c_i^2 < \varepsilon.$$

This shows that $(v_n)_{n=0}^{\infty}$ is Cauchy, so by the completeness of V, $v_n \to w$ for some $w \in V$. It remains to show that $v = w$.

But by using the continuity of $\langle \cdot, \cdot \rangle$ at each argument,

$$\langle w, \xi_i \rangle = \left\langle \lim_{n\to\infty} v_n, \xi_i \right\rangle = \lim_{n\to\infty} \langle v_n, \xi_i \rangle = c_i = \langle v, \xi_i \rangle.$$

This means w, v have the same sequence of $\mathcal{X}$-coefficients, so by Proposition 4.23, $w = v$, as required.

QED

The statement of (b) of Proposition 4.25 could equivalently be phrased as saying that $\lim_{n\to\infty} v_n = v$, where the limit is considered in the metric defined by the norm $\|\cdot\|$. Since the v_n's are precisely the finite partial sums of the series $\sum_{i=0}^{\infty} c_i \xi_i$, one can thus say that Proposition 4.25 provides us with a representation of v as the sum of a convergent series in V:

$$v = \sum_{i=0}^{\infty} c_i \xi_i.$$

This way of phrasing Proposition 4.25 is known as the *Riesz-Fischer theorem*.

A useful consequence of the Riesz-Fischer theorem is the following statement about norms of vectors in our Hilbert space V.

Proposition 4.26. Parseval's Identity

Let $v \in V$ and let $(c_i)_{i=0}^{\infty}$ be the $\mathcal{X}$-coefficients of v. Then

$$\|v\|^2 = \sum_{i=0}^{\infty} c_i^2.$$

Proof. For all $n \in \mathbb{N} \cup \{0\}$, let $v_n \in V_n$ be defined as $v_n = \sum_{i=0}^{n} c_i \xi_i$. Then by Proposition 4.25, we have $\lim_{n\to\infty} v_n = v$. It follows that $\lim_{n\to\infty} \|v_n\| = \|v\|$. Thus,

$$\|v\|^2 = \lim_{n\to\infty} \|v_n\|^2 = \lim_{n\to\infty} \sum_{i=0}^{n} c_i^2 = \sum_{i=0}^{\infty} c_i^2.$$

QED

The next corollary records a consequence of Parseval's identity which is related to a tatement known as the *Riemann-Lebesgue theorem*. We will see shortly the Riemann-Lebesggue theorem stated precisely in the setting of functions on the interval $[-\pi, \pi]$.

Corollary 4.26.1.

Let $v \in V$ and let $(c_i)_{i=0}^{\infty}$ be the $\mathcal{X}$-coefficients of v. Then $\lim_{i\to\infty} c_i = 0$.

Quite Clear!

9. $\mathrm{L}^2\left(\lambda_{[-\pi,\pi]}\right)$ and Its Natural Orthonormal Basis

The goal of this subsection is to see what the results from the previous subsection tells us about hte special Hilbert space

$$V = \mathrm{L}^2\left(\lambda_{[-\pi,\pi]}\right).$$

In order to invoke what we did in the previous subsection, we first need to clarify what is the orthonormal basis of $\mathrm{L}^2\left(\lambda_{[-\pi,\pi]}\right)$ that we want to work with.

Notation 4.21. φ_n, ξ_n

We define functions $\varphi_0, \varphi_1, \ldots : [-\pi, \pi] \to \mathbb{R}$ in the way described as follows.

First we let $\varphi_0 : [-\pi, \pi] \to \mathbb{R}$ be the function identically equal to $\frac{1}{\sqrt{2\pi}}$.

For every $k \in \mathbb{N}$, we define $\varphi_{2k-1}, \varphi_{2k}$ by putting

$$\varphi_{2k-1}(t) = \frac{1}{\sqrt{\pi}} \sin(kt)$$
$$\varphi_{2k}(t) = \frac{1}{\sqrt{\pi}} \cos(kt)$$

for all $t \in [-\pi, \pi]$. It is clear that for every $n \in \mathbb{N} \cup \{0\}$, we have $\varphi_n \in C([-\pi, \pi], \mathbb{R})$, which implies in particular that $\varphi_n \in \mathcal{L}^2\left(\lambda_{[-\pi,\pi]}\right)$. It thus makes sense to put

$$\xi_n = \hat{\varphi_n}$$

for all $n \in \mathbb{N}$, where $\hat{\varphi_n}$ denotes the coset $\varphi_n + \mathcal{N}$ ($\mathcal{N}$ is the null space of the seminorm on $\mathcal{L}^2\left(\lambda_{[-\pi,\pi]}\right)$).

The family of vectors $\{\xi_n\}_{n=0}^{\infty}$ is our candidate of orthonormal basis for the space $\mathrm{L}^2\left(\lambda_{[-\pi,\pi]}\right)$.

For each even $n \in \mathbb{N} \cup \{0\}$, φ_n is an even function (i.e. $\varphi_n(-t) = \varphi_n(t)$ for all $t \in [-\pi, \pi]$). On the other hand, for an odd $n \in \mathbb{N} \cup \{0\}$, φ_n is an odd function (i.e. $\varphi_n(-t) = -\varphi_n(t)$ for all $t \in [-\pi, \pi]$).

The normalization constant $\frac{1}{\sqrt{\pi}}$ in the definitions of φ_n's have the role to ensure that

$$\int_{[-\pi,\pi]} \varphi_n^2 \, \mathrm{d}\lambda_{[-\pi,\pi]} = 1 \qquad [4.27]$$

for all $n \in \mathbb{N} \cup \{0\}$. This can be checked by using Riemann integrals.

When putting on *hats*, [4.27] implies

$$\|\xi_n\| = 1$$

for all $n \in \mathbb{N} \cup \{0\}$.

We can also prove the following.

Exercise 4.7.

Prove that for all $m \neq n$ in $\mathbb{N} \cup \{0\}$, we have

$$\int_{-\pi,\pi} \varphi_m \varphi_n \, \mathrm{d}\lambda_{[-\pi,\pi]} = 0.$$

From above, we infer that $\{\xi_n\}_{n=0}^{\infty}$ is an orthonormal subset of V. Our next job is to prove that cl-span $\{\xi_n\}_{n=1}^{\infty} = V$. This will come out as an application of the Stone-Weierstrass theorem, which enters the game by way of the following two lemmas.

Lemma 4.27.

Let

$$\mathcal{T} = \mathrm{span}\left\{\varphi_n\right\}_{n=0}^{\infty} \subseteq C([-\pi, \pi], \mathbb{R}). \qquad [4.28]$$

Then $\mathcal{T}$ is closed under multiplication, and is therefore an algebra of continuous functions on $[-\pi, \pi]$.

Use Trigonometric Identities

$C([-\pi,\pi], \mathbb{R})$ is a Banach space when considered with the norm $\|\cdot\|_\infty$ defined by

$$\|f\|_\infty = \sup_{t\in[-\pi,\pi]} |f(t)|$$

for all $f \in C([-\pi,\pi], \mathbb{R})$. We chase the preceding lemma with a statement about the $\|\cdot\|_\infty$-closure of $\mathcal{T}$.

Lemma 4.28.
Let $\mathcal{T}$ be defined as in [4.28]. Then

$$\|\cdot\|_\infty - \mathrm{cl}(\mathcal{T}) = \{\varphi \in C([-\pi,\pi], \mathbb{R}) : \varphi(-\pi) = \varphi(\pi)\}.$$

Use Stone-Weierstrass

Having made the above preparations, we can now claim that our choice of ξ_n's produces indeed an orthonormal basis for the space $\mathrm{L}^2(\lambda_{[-\pi,\pi]})$.

Proposition 4.29.
cl-span $\{\xi_n\}_{n=0}^\infty = V$.

10. Fourier Series in $\mathrm{L}^2(\lambda_{[-\pi,\pi]})$

In this subsection we enjoy the benefits of having pinned down an orthonormal basis for the L^2-space under consideration. We have seen that it is very useful that every vector in the space we consider what we called *$\mathcal{X}$-coefficients* of that vector, with an orthonormal basis $\mathcal{X} = \{\xi_i\}_{i=0}^\infty$ for the space. In the case at hand, these $\mathcal{X}$-coefficients are precisely what one calls *Fourier coefficients*. Their explicit definition is recorded below. Note that in the next definition there is no problem to introduce Fourier coefficients for an actual function $f \in \mathcal{L}^2(\lambda_{[-\pi,\pi]})$.

Def'n 4.22. **Fourier Coefficient** of a Function
Let $f \in \mathcal{L}^2(\lambda_{[-\pi,\pi]})$. For all $n \in \mathbb{N} \cup \{0\}$, we call

$$c_n = \langle f, \varphi_n \rangle = \int_{[-\pi,\pi]} f\varphi_n \,\mathrm{d}\lambda_{[-\pi,\pi]}$$

the *N-**th Fourier coefficient*** of f.

When spelled out explicitly, the Fourier coefficients of f are thus described as follows. First, we have

$$c_0 = \frac{1}{\sqrt{2\pi}} \int_{[-\pi,\pi]} f \,\mathrm{d}\lambda_{[-\pi,\pi]}.$$

Then for every $k \in \mathbb{N}$, we have

$$c_{2k-1} = \frac{1}{\sqrt{\pi}} \int_{[-\pi,\pi]} f(t)\sin(kt)\,\mathrm{d}\lambda_{[-\pi,\pi]}(t),$$
$$c_{2k} = \frac{1}{\sqrt{\pi}} \int_{[-\pi,\pi]} f(t)\cos(kt)\,\mathrm{d}\lambda_{[-\pi,\pi]}(t).$$

Let us then restate, in our present terminology, the two main results from Subsection 8: the Riesz-Fischer theorem and Parseval's identity.

Proposition 4.30. Riesz-Fischer Theorem
Let $f \in \mathcal{L}^2(\lambda_{[-\pi,\pi]})$ and let $(c_i)_{i=0}^\infty$ be the Fourier coefficients of f. For every $n \in \mathbb{N} \cup \{0\}$ consider the trigonometric polynomial

$$p_n = \sum_{i=0}^n c_i \varphi_i.$$

Then $\lim_{n\to\infty} \|p_n - f\|_2 = 0$.

Quite Clear!

Same as we discussed in Subsection 8, the conclusion of the Riesz-Fischer theorem can be viewed as a representation of f in the form of a $\|\cdot\|_2$-convergent series,

$$f = \sum_{i=0}^{\infty} c_i \varphi_i.$$

The series on the right-hand side goes under the name of *Fourier series* of f.

Concerning the description of the $\|\cdot\|_2$-norm of f in terms of its Fourier coefficients, Parseval's identity now takes the following form.

Proposition 4.31. Parseval's Identity

Let $f \in \mathcal{L}^2\left(\lambda_{[-\pi,\pi]}\right)$ and let $(c_i)_{i=0}^{\infty}$ be the Fourier coefficients of f. Then

$$\|f\|_2^2 = \sum_{i=0}^{\infty} c_i^2.$$

Quite Clear!

Corollary 4.31.1.

Let $f \in \mathcal{L}^2\left(\lambda_{[-\pi,\pi]}\right)$ and let $(c_i)_{i=0}^{\infty}$ be the Fourier coefficients of f. Then

$$\lim_{i\to\infty} c_i = 0.$$

11. Fourier Coefficients for $\mathcal{L}^1$-functions

Let $f \in \mathcal{L}^1\left(\lambda_{[}\,[-\pi,\pi]]\right)$ and let $n \in \mathbb{N} \cup \{0\}$. Then it makes sense to consider

$$c_n = \int_{[-\pi,\pi]} f\varphi_n \, \mathrm{d}\lambda_{[-\pi,\pi]}. \tag{4.29}$$

Def'n 4.23.

Let $f \in \mathcal{L}^1\left(\lambda_{[-\pi,\pi]}\right)$ and let $n \in \mathbb{N} \cup \{0\}$. We call c_n in [4.29] the ***n-th Fourier coefficient*** of f.

We have estimates on c_n given by

$$|c_n| \le \frac{1}{\sqrt{\pi}} \|f\|_1, \tag{4.30}$$

which comes directly from the definition of φ_n's.

First, observe that

$$|f\varphi_n| \le \frac{1}{\sqrt{\pi}} |f|,$$

so that

$$\int_{[-\pi,\pi]} |f\varphi_n| \, \mathrm{d}\lambda_{[-\pi,\pi]} \le \int_{[-\pi,\pi]} \frac{1}{\sqrt{\pi}} |f| \, \mathrm{d}\lambda_{[-\pi,\pi]} = \frac{1}{\sqrt{\pi}} \|f\|_1,$$

showing $f\varphi_n \in \mathcal{L}^1\left(\lambda_{[-\pi,\pi]}\right)$, which justifies the definition [4.29].

Furthermore,

$$|c_n| = \left| \int_{[-\pi,\pi]} f\varphi_n \, \mathrm{d}\lambda_{[-\pi,\pi]} \right| \le \int_{[-\pi,\pi]} |f\varphi_n| \, \mathrm{d}\lambda_{[-\pi,\pi]} \le \frac{1}{\sqrt{\pi}} \|f\|_1,$$

yielding [4.30].

The above discussion is relevant to Fourier analysis for $\mathcal{L}^1$-functions, since $\mathcal{L}^1\left(\lambda_{[-\pi,\pi]}\right)$ properly contains $\mathcal{L}^2\left(\lambda_{[-\pi,\pi]}\right)$. Fourier analysis for $\mathcal{L}^1$-functions turns out to be more difficult than for $\mathcal{L}^2$. We can nevertheless establish a property which continues to work in $\mathcal{L}^1$-framework – *Riemann-Lebesgue theorem*.

Notation 4.24. $\underline{c}_n$
For each $n \in \mathbb{N} \cup \{0\}$, let $\underline{c}_n : \mathcal{L}^1\left(\lambda_{[-\pi,\pi]}\right) \to \mathbb{R}$ be defined by

$$\underline{c}_n(f) = n\text{-th Fourier coefficient of } f = \int_{[-\pi,\pi]} f\varphi_n \, \mathrm{d}\lambda_{[-\pi,\pi]}.$$

Observe that each $\underline{c}_n$ is a linear functional:

$$\underline{c}_n(af+g) = \int_{[-\pi,\pi]} (af+g)\,\varphi_n \, \mathrm{d}\lambda_{[-\pi,\pi]} = a\int_{[-\pi,\pi]} f\varphi_n \, \mathrm{d}\lambda_{[-\pi,\pi]} + \int_{[-\pi,\pi]} g\varphi_n \, \mathrm{d}\lambda_{[-\pi,\pi]} = a\underline{c}_n(f) + \underline{c}_n(g).$$

We also have

$$|\underline{c}_n(f)| \le \frac{1}{\sqrt{\pi}} \|f\|_1$$

for all $f \in \mathcal{L}^1\left(\lambda_{[-\pi,\pi]}\right)$ by [4.30]. This means $\underline{c}_n$ is bounded. Combining linearity and boundedness, we obtain that

$$|\underline{c}_n(f) - \underline{c}_n(g)| \le \frac{1}{\sqrt{\pi}} \|f-g\|_1$$

for all $f, g \in \mathcal{L}^1\left(\lambda_{[-\pi,\pi]}\right)$, showing that $\underline{c}_n$ is $\frac{1}{\sqrt{\pi}}$-Lipschitz.

For any $f \in \mathcal{L}^2\left(\lambda_{[-\pi,\pi]}\right)$, we know that (Corollary 4.31.1)

$$\lim_{n\to\infty} \underline{c}_n(f) = 0$$

as a consequence of Parseval's identity. This contiues to hold for any $f \in \mathcal{L}^1\left(\lambda_{[-\pi,\pi]}\right)$.

Theorem 4.32. Riemann-Lebesgue Theorem

Let $f \in \mathcal{L}^1\left(\lambda_{[-\pi,\pi]}\right)$. Then

$$\lim_{n\to\infty} \underline{c}_n(f) = 0. \tag{4.31}$$

Proof. Fix $f \in \mathcal{L}^1\left(\lambda_{[-\pi,\pi]}\right)$, for which we will verify [4.31] holds. Let $\varepsilon > 0$ be given. We have to find $n_0 \in \mathbb{N}$ such that

$$\textit{for all } n \ge n_0 \textit{ we have } |\underline{c}_n(f)| < \varepsilon. \tag{4.32}$$

Since $C([-\pi,\pi],\mathbb{R})$ is dense in $\mathcal{L}^1\left(\lambda_{[-\pi,\pi]}\right)$, there is $g \in C([-\pi,\pi],\mathbb{R})$ such that

$$\|f-g\|_1 < \frac{\varepsilon}{2}.$$

Moreover, sicne $g \in C([-\pi,\pi],\mathbb{R}) \subseteq \mathcal{L}^2\left(\lambda_{[-\pi,\pi]}\right)$, we have that

$$\lim_{n\to\infty} \underline{c}_n(g) = 0.$$

Hence we have $n_0 \in \mathbb{N}$ such that

$$|\underline{c}_n(g)| < \frac{\varepsilon}{2}$$

for all $n \ge n_0$.

We now claim that this particular n_0 satisfies [4.32]: given any $n \ge n_0$

$$|\underline{c}_n(f)| = |(\underline{c}_n(f) - \underline{c}_n(g)) + \underline{c}_n(g)| \le |\underline{c}_n(f) - \underline{c}_n(g)| + |\underline{c}_n(g)| \le \frac{1}{\sqrt{\pi}} \|f-g\|_1 + \frac{\varepsilon}{2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

QED

In this sectino, we look at a continuous function on $[-\pi, \pi]$: let

$$f \in C([-\pi, \pi], \mathbb{R}).$$

Since f is continuous, $f \in \mathcal{L}^1(\lambda_{[-\pi,\pi]})$, so has Fourier coefficients:

$$c_i = \int_{[-\pi,\pi]} f\varphi_i \, \mathrm{d}\lambda_{[-\pi,\pi]} = \int_{-\pi}^{\pi} f(t)\,\varphi_i(t)\,\mathrm{d}t,$$

where the last equality follows from the fact that f is continuous.

Consider the trigonometric polynomials

$$\psi_n = \sum_{i=0}^{n} c_i\varphi_i$$

for all $n \in \mathbb{N} \cup \{0\}$. Since $f \in \mathcal{L}^2(\lambda_{[-\pi,\pi]})$ in particular, the Riesz-Fischer theorem tells us that $\psi_n \to f$ as $n \to \infty$ with respect to $\|\cdot\|_2$. That is,

$$\lim_{n\to\infty} \|\psi_n - f\|_2 = 0, \tag{4.33}$$

which can be also written as

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} (\psi_n(t) - f(t))^2 \,\mathrm{d}t = 0.$$

Since f and every ψ_n's are in $C([-\pi, \pi], \mathbb{R})$, where we like to use the supremum norm $\|\cdot\|_\infty$, we ask:

couldn't it be true that we actually have $\psi_n \to f$ as $n \to \infty$ with respect to $\|\cdot\|_\infty$? [4.34]

Note that [4.34] is stronger than [4.33]. Unfortunately, the answer for [4.34] is negative, but counterexamples are subtle.

A surprising twist: let us average the ψ_n's, and put

$$\sigma_n = \frac{1}{n+1}(\psi_0 + \cdots + \psi_n) = \frac{1}{n+1}\sum_{i=0}^{n}\psi_i, \qquad \forall n \in \mathbb{N} \cup \{0\}. \tag{4.35}$$

Note that each σ_n is still a trigonometric polynomial. By a direct calculation, we can verify that

$$\sigma_n = \sum_{i=0}^{n} c_i \frac{n+1-i}{n+1}\psi_i, \ \forall n \in \mathbb{N} \cup \{0\}. \tag{4.36}$$

Then Fejer's theorem says that σ_n's do the job.

Theorem 4.33. Fejer's Theorem

Consider $(\sigma_n)_{n=0}^{\infty}$ defined above. Then

$$\sigma_n \to f$$

with respect to $\|\cdot\|_\infty$ as $n \to \infty$.

The important part of the proof of Fejer's theorem lies in the following fact: for all $n \in \mathbb{N} \cup \{0\}$, there is a continuous function K_n of two variables such that

$$\psi_n(t) = \int_{-\pi}^{\pi} f(s)\,K_n(s,t)\,\mathrm{d}s$$

for all $t \in [-\pi, \pi]$. For instance,

$$\begin{aligned}
\psi_2(t) &= c_0\varphi_0(t) + c_1\varphi_1(t) + c_2\varphi_2(t) \\
&= \left(\int_{-\pi}^{\pi} f(s)\,\varphi_0(s)\,\mathrm{d}s\right)\varphi_0(t) + \left(\int_{-\pi}^{\pi} f(s)\,\varphi_1(s)\,\mathrm{d}s\right)\varphi_1(t) + \left(\int_{-\pi}^{\pi} f(s)\,\varphi_2(s)\,\mathrm{d}s\right)\varphi_2(t) \\
&= \int_{-\pi}^{\pi} f(s)\sum_{n=0}^{2}\varphi_n(s)\,\varphi_n(t)\,\mathrm{d}s.
\end{aligned}$$

This means

$$K_n(s,t) = \sum_{n=0}^{2} \varphi_n(s)\, \varphi_n(t)\,.$$

In general, we obtain

$$K_n(s,t) = \sum_{i=0}^{n} \varphi_i(s)\, \varphi_i(t)\,, \qquad \forall n \in \mathbb{N} \cup \{0\}\,, \tag{4.37}$$

on which we can use trigonometric identities to obtain an explicit formula for $K_n(s,t)$, leading to a function called *Dirichlet kernel*.

When we average $\sigma_n = \frac{1}{n+1} \sum_{i=0}^{n} \psi_i$, we get

$$\sigma_n(t) = \int_{-\pi}^{\pi} f(s)\, F_n(s,t)\ \mathrm{d}s,$$

where $F_n(s,t) = \sum_{i=0}^{n} K_n(s,t)$. These F_n's lead to a function called the *Fejer kernel*.

With these kernels in hand, Fejer's theorem becomes:

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} f(s)\, F_n(s,t)\ \mathrm{d}s - f(t) = 0$$

uniformly for all $t \in [-\pi, \pi]$.